

AC AAR79163;
 XX
 DT 22-DEC-1995 (first entry)
 XX
 DE Partial sequence of bovine alpha 3 chain of type IV collagen.
 XX
 KW Type IV collagen; alpha 3 chain.
 XX
 OS Bos taurus.
 XX
 PN US5424408-A.
 XX
 PD 13-JUN-1995.
 XX
 PR 30-NOV-1990; 90US-0621091.
 XX
 PR 30-NOV-1990; 90US-0621091.
 XX
 (UNIV) UNIV KANSAS MEDICAL CENT.
 (UYVA) UNIV YALE.
 PI Hudson BG, Morrison KE, Reeders ST;
 XX
 DR WPI: 1995-262631/34.
 XX
 N-PSDB; AAQ96290.
 XX
 cDNAs encoding human or bovine alpha-3 type 4 collagen peptide(s) -
 PT useful for detection and therapeutic removal of antibodies associated
 PT with Goodpasture syndrome
 XX
 Disclosure; Columns 5-8; 33pp; English.
 XX
 CC Using the PCR with primers derived from each end of the known 27 AA
 CC residue bovine alpha 3 (IV) collagen protein sequence, a 68 bp
 CC bovine genomic fragment was amplified. This fragment was then used
 CC to a bovine lens cDNA library and a 1.5 kb partial cDNA clone was
 CC obt'd. (clone KMC15). This encodes 238 residues of the triple helical
 CC collagenous domain and all 233 residues of the C terminal non-
 CC collagenous domain of the alpha 3 (IV) chain. An isolated and
 CC substantially pure nt. having the sequence in AAQ96290 is claimed.
 XX
 Sequence 471 AA:
 SQ

Query Match 93.7%; Score 986; DB 16; Length 471;
 Best Local Similarity 92.6%; Pred. No. 6.2e-103; Gaps 0;
 Matches 176; Conservative 10; Mismatches 4; Indels 0;
 Gaps 0;

1	ORAHGDLGTIGSCLQRFITMPFLCNVNDVCNFRNDYSWJLSTPALMPMMAPITGR	60
:	: : : : : : : : : :	
Db	eqahqdgltiqrltqslqrftmpflcnindvnfasrndyswistpamipdmmapitgr	340
Qy	61 ALEPTISRCTVEGEGPAIAAVHSQTDIIPCPHGWISLNWGFSFTMFTSGSEGTQALA	120
:	: : : : : : : : : :	
Db	341 amepyisrctveoqpaiaavhsqtdipcpagwislwkqfimftagsseggala	400
Qy	121 SFGSCLEEFRASPFLECHGRGTCNYNSNSFWLNLNPERRKPKIPSPVKABLEKII	180
Db	401 spgscleefraspflechgrgtnyynsnsfswlslpkrmfrkipstvkgelenii	460
Qy	181 SRCQYCMKKR 190	
Db	461 srcqvcmkmr 470	

Search completed: March 6, 2002, 06:56:24
 Job time: 135 sec

KW	type IV collagen alpha chain monomer; human; inhibitor; angiogenesis; tumour growth; integrin receptor; carcinoma; rhabdomyosarcoma; retinoblastoma; Ewing sarcoma; neuroblastoma; osteosarcoma; Leukaemia; diabetic retinopathy; rheumatoid arthritis; neovascularisation; muscular degeneration; corneal graft rejection; vitamin A deficiency; atopic keratitis; Mycobacteria infection; chemical burn; sarcoid; Kaposi's sarcoma; sickle cell anaemia; carotid obstructive disease; chronic inflammation; psoriasis; therapy; alpha3(IV)NC1.
KW	Homo sapiens.
XX	
PN	WO20059532-A1.
XX	
PD	12-OCT-2000.
XX	
PF	31-MAR-2000; 2000WO-US08678.
XX	
PR	01-APR-1999; 99US-0127391.
XX	
(BIOS-) BIOSTRATUM INC.	
PI	Brooks P, Hudson B;
XX	
WPI:	2000-664962/64.
DR	N-PSDB; AAQ96993.
XX	
PT	Use of antagonists of specific integrin receptors for inhibiting angiogenesis, tumour growth or metastases, or endothelial cell interactions with the extracellular matrix
PT	Disclosure; Fig 17c; 78pp; English.
XX	
CC	This sequence is a human type IV collagen alpha chain monomer, designated alpha3(IV)NC1. The invention relates to a method for inhibiting angiogenesis, tumour growth or metastases, or endothelial cell interactions with the extracellular matrix, comprising contacting the cells or tissue with a polypeptide composition containing antagonists of specific integrin receptors. The methods and the antagonists are useful for inhibiting angiogenesis, tumour growth or metastases, or endothelial cell interaction with the extracellular matrix. The antagonists are also useful for treating diseases and conditions with accompanying undesired angiogenesis, e.g. solid and blood-borne tumours (e.g. melanomas, carcinomas, sarcomas, rhabdomyosarcoma, retinoblastoma, Ewing sarcoma, neuroblastoma, osteosarcoma or leukaemia). These are also applicable to treating non-tumorigenic diseases and conditions with accompanying undesired angiogenesis, e.g. diabetic retinopathy, rheumatoid arthritis, retinal neovascularisation, choroidal neovascularisation, muscular degeneration, corneal graft rejection, vitamin A deficiency, atopic keratitis, Mycobacteria infections, chemical burns, Kaposi's sarcoma, sickle cell anaemia, sarcoid, carotid obstructive disease, post-laser complications, chronic inflammation or psoriasis.
CC	SQ Sequence 268 AA;
CC	Query Match 99.5%; Score 1047; DB 21; Length 268; Best Local Similarity 93.5%; Pred. No. 3.7e-110; Matches 190; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY	1 QRAHQDQDIGTLGSLCIQRRTMPFLFCNNNDVCNEASRDNYSYNLSPTALPMNMNAPITGR 60
Db	78 qrahqdgltgslcqrftmpflfcnnndvcnfasrdnysylstpalpmnmnmapitgr 137
QY	61 ALEYVYISRCRVCSGPALIAVHSQTTDIPCPKGWISWKGFERTIMFSSAGSGCTGQAL 120
Db	138 alepyisrcvcegpalalavnsqtdipcpchgwlsiwkgfsimftsagsagggala 197
QY	121 SPSCLEERFRASPLECHRGTCNYYSYSFWLASLPERMRKPITSTVKAGELEXI 180
Db	198 spgscleerfraspflechrgtcnyysysfwlaslpermrkpitstvkageleki 257
QY	181 SROVCMKRR 191
XX	
CC	Sequence 218 AA;
CC	SQ Sequence 218 AA;
CC	Query Match 97.4%; Score 1025; DB 16; Length 218; Best Local Similarity 97.9%; Pred. No. 8.6e-108; Matches 187; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY	1 QRAHQDQDIGTLGSLCIQRRTMPFLFCNNNDVCNEASRDNYSYNLSPTALPMNMNAPITGR 60
Db	28 qrahqdgltgslcqrftmpflfcnnndvcnfasrdnysylstpalpmnmnmapitgr 87
QY	61 ALEYVYISRCRVCSGPALIAVHSQTTDIPCPKGWISWKGFERTIMFSSAGSGCTGQAL 120
Db	88 alepyisrcvcegpalalavnsqtdipcpchgwlsiwkgfsimftsagsagggala 147
QY	121 SPSCLEERFRASPLECHRGTCNYYSYSFWLASLPERMRKPITSTVKAGELEXI 180

RESULT	4	CC	This sequence represents a recombinant type IV collagen
AAV31993		CC	non-collagenous (NC1) domain alpha-3 polypeptide composed of a
ID		CC	BM40 signal sequence (which is cleaved from the mature protein) to
XX		CC	facilitate protein secretion, and a mature protein comprising an
AC		CC	affinity tag (facilitates purification and identification of the
XX		CC	material) and the alpha-1 chain monomer. The invention provides
XX		CC	methods and kits for inhibiting angiogenesis, tumour growth and
DT		CC	metastasis, and endothelial cell interaction with the extracellular
DE		CC	tissue with 1 or more isolated type IV collagen NC1 alpha chain
XX		CC	monomer(s) selected from the group consisting of alpha-1, alpha-2,
XX		CC	alpha-3 and alpha-6 NC1 chain monomers (see AAV31991-96). The
XX		CC	monomers can be produced via recombinant protein expression. The
XX		CC	polynucleotides and polypeptides are used to treat an angiogenesis-
XX		CC	mediated disorder or condition, especially selected from solid and
XX		CC	blood-borne tumours, diabetic retinopathy, rheumatoid arthritis,
XX		CC	retinal neovascularization, choroidal neovascularization, macular
XX		CC	degeneration, corneal neovascularization, retinopathy of prematurity,
XX		CC	corneal graft rejection, neovascular glaucoma, retrobulbar
XX		CC	fibroplasia, epidemic keratoconjunctivitis, vitamin A deficiency,
XX		CC	contact lens overwear, atopic keratitis, superior limbic keratitis,
XX		CC	superior limbic keratitis, sicca, soggens, acne rosacea, phlyctenulosis,
XX		CC	syphilis, mycobacteria infections, lipid degeneration, chemical
XX		CC	burns, bacterial ulcers, fungal ulcers, herpes simplex infections,
XX		CC	herpes zoster infections, protozoan infections, kaposi's sarcoma,
XX		CC	Mooren ulcer, Terrien's marginal degeneration, marginal keratolysis,
XX		CC	trauma, eye wear, atopic keratitis, superior limbic keratitis,
XX		CC	scleritis, Steven's Johnson disease, radial keratotomy, sickle cell
XX		CC	anaemia, sarcoid, pseudoxanthoma elasticum, Paget's disease, vein
XX		CC	occlusion, artery occlusion, carotid obstructive disease, chronic
XX		CC	uveitis, chronic vitritis, Lyme's disease, Eales disease, Bechets
XX		CC	disease, myopia, optic pits, Stargardt's disease, pars planitis,
XX		CC	chronic retinal detachment, hyperviscosity syndromes, toxoplasmosis,
XX		CC	post-laser complications, abnormal proliferation of fibrovascular
XX		CC	tissue, haemangioma, Osler-Weber-Rendu, AIDS, ocular neovascular
XX		CC	disease, osteoarthritis, chronic inflammation, Crohn's disease,
XX		CC	ulcerative colitis, psoriasis, atherosclerosis, and pemphigoid (all
OS		CC	claimed).
OS		SQ	Sequence 268 AA:
XX			
FH	Key		location/Qualifiers
FT	peptide	1..17	"BM40 signal peptide"
FT	Protein	18..268	/note= "mature protein"
FT	peptide	18..25	/note= "affinity tag"
FT	Protein	26..268	/note= "NC1 alpha-3 monomer"
XX	W09949885-A2.		
XX	07-OCT-1999.		
XX	26-MAR-1999;	99WO-US06445.	
XX	27-MAR-1998;	98US-0079783.	
PR	29-OCT-1998;	98US-0106170.	
PA	(UNIV) UNIV KANSAS MEDICAL CENT.		
XX	Hudson BG, Sarras MP;		
PT	WPI; 1999-601297/51.		
DR	N-PSDB; AA20091.		
XX	Inhibition of angiogenesis with non-collagenous alpha chain monomer		
PT	useful for treating e.g. tumor growth or metastasis,		
PT	neovascularisation, etc		
XX	Disclosure; Fig 17c; 56pp; English.		
RESULT	5	CC	
XX	AAV97555	CC	
ID	AAV97555 standard; Protein; 268 AA.	CC	
XX		CC	
AC	AAV97555;	CC	
XX	12-FEB-2001 (first entry)	CC	
DE	Human alpha3(IV)NC1 protein sequence.	CC	

Db	148	sgscleefraspflechrgtccnyysnsyswlasinpermfrkipstvkageleki	207
Db	115	alepyisrctvcegpaiaavinavsqtdipcpghwislwgfisimtsaysegqala	174
Qy	121	SPGSCLEEFRAFPFLECHRGRCNYYSNSYSWLASINPERMFRKIPSTVKAGELEKII	180
Db	175	SPGSCLEEFRAFPFLECHRGRCNYYSNSYSWLASINPERMFRKIPSTVKAGELEKII	234
Qy	181	SPCQVCMKKRH 191	
Qy	235	srcqvcmkkrh 245	
RESULT	2		
ID	AAY44172	standard; Protein; 218 AA.	
XX			
AC	AAY44172;		
XX			
DT	01-FEB-2000	(first entry)	
Human type IV collagen alpha3 chain protein.			
XX			
KW	Recombinant; bovine; alpha3 chain; type IV collagen; detection; Goodpasture syndrome; antibody; blood; tissue; human; nephrotrophism.		
XX			
OS	Homo sapiens.		
XX			
US	US5973120-A.		
XX			
PD	26-OCT-1999.		
XX			
PF	07-MAR-1995;	95US-0399889.	
XX			
PR	30-NOV-1990;	90US-0621091.	
XX			
PA	(UYYA) UNIV YALE.		
PA	(UNIV) UNIV KANSAS MEDICAL CENT.		
XX			
PI	Hudson BG, Reeder ST, Morrison KE,		
XX			
DR	WPI; 1999-610317/52.		
DR	N-PSDB; AAZ28775.		
XX			
PT	Isolated alpha 3 chain of type IV collagen polypeptide useful for diagnosis and treatment of Goodpasture syndrome		
XX			
PS	Claim 2; Column 35-36; 27pp; English.		
CC	This sequence represents a recombinant human alpha3 chain of type IV collagen polypeptide. The sequence corresponds to the 218 amino acids of the C-terminal non-collagenous domain. Alpha3 chain collagen polypeptides are useful for detecting Goodpasture antibodies in blood or tissue from a human patient and for treating Goodpasture syndrome, especially by neutralising the antibodies in the blood. The polypeptides also have a nephrotrophic activity.		
CC	XX		
SQ	Sequence 218 AA;		
Query Match	99.5%	Score 1047; DB 20; Length 218;	
Best Local Similarity	99.5%	Pred. No. 2.8e-110; Matches 190; Conservative 0; Mismatches 1; Indels 0; Gaps 0;	
Qy	1	ORAHGQDGLTGLGSCLQRFTTMPLFCNNNDVNFAASRNDYSWLSTPALMPNMAPTRG 60	
Db	28	qrahggdgltglgsclqrfttmplfcvndvenfasrndyswlstpalmppnnatgr 87	
Qy	61	ALEPYISRCTVCEGPAIAAVHSQTDDIPCPCHGWISLWGFSIMTSAGSEGTOALA 120	
Db	88	alepyisrctvcegpaiaavinavsqtdipcpghwislwgfisimtsaysegqala 147	
Qy	121	SPGSCLEEFRAFPFLECHRGRCNYYSNSYSWLASINPERMFRKIPSTVKAGELEKII 180	
Db	148	SPGSCLEEFRAFPFLECHRGRCNYYSNSYSWLASINPERMFRKIPSTVKAGELEKII 207	
Qy	181	SPCQVCMKKRH 191	
Db	208	srcqvcmkkrh 218	

GenCore version 4.5
copyright (c) 1993 - 2000 Compugen Ltd.

ALIGNMENTS

OM protein - protein search, using sw model
Run on: March 6, 2002, 06:54:09 ; Search time 54.67 Seconds
(without alignments)
Scoring table: BLOSUM62
Gappen 10.0 , Gapext 0.5

Title: US-09-543-371-10_COPY_54_244

Perfect score: 1052

Sequence: 1 QRAHGDLGTLGSLCLQRFTT..... KAGELEKILSRCQVCMKKRH 191

Scoring table: 258.789 Million cell updates/sec

Searched: 522463 seqs, 74073290 residues

To number of hits satisfying chosen parameters: 9

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 90%
Maximum Match 100%
Listing first 1000 summaries

Database : A_Geneseq
1: /SIDS8/geodata/geneseq/geneseq/AA1980.DAT:
2: /SIDS8/geodata/geneseq/geneseq/AA1981.DAT:
3: /SIDS8/geodata/geneseq/geneseq/AA1982.DAT:
4: /SIDS8/geodata/geneseq/geneseq/AA1983.DAT:
5: /SIDS8/geodata/geneseq/geneseq/AA1985.DAT:
6: /SIDS8/geodata/geneseq/geneseq/AA1986.DAT:
7: /SIDS8/geodata/geneseq/geneseq/AA1987.DAT:
8: /SIDS8/geodata/geneseq/geneseq/AA1988.DAT:
10: /SIDS8/geodata/geneseq/geneseq/AA1989.DAT:
11: /SIDS8/geodata/geneseq/geneseq/AA1990.DAT:
12: /SIDS8/geodata/geneseq/geneseq/AA1991.DAT:
13: /SIDS8/geodata/geneseq/geneseq/AA1992.DAT:
14: /SIDS8/geodata/geneseq/geneseq/AA1993.DAT:
15: /SIDS8/geodata/geneseq/geneseq/AA1994.DAT:
16: /SIDS8/geodata/geneseq/geneseq/AA1995.DAT:
17: /SIDS8/geodata/geneseq/geneseq/AA1996.DAT:
18: /SIDS8/geodata/geneseq/geneseq/AA1997.DAT:
19: /SIDS8/geodata/geneseq/geneseq/AA1998.DAT:
20: /SIDS8/geodata/geneseq/geneseq/AA2000.DAT:
21: /SIDS8/geodata/geneseq/geneseq/AA2001.DAT:
22: /SIDS8/geodata/geneseq/geneseq/AA2001.DAT:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	1052	100.0	245 21	AAY67942 Human type IV collagen
2	1047	99.5	218 20	AAY44172 Human type IV collagen
3	1047	99.5	218 21	AAY56784 Human alpha3 type IV collagen N
4	1047	99.5	268 20	AAY31993 Human alpha3(V)NC
5	1047	99.5	268 21	AAY97555 Partial sequence o
6	1025	97.4	16 16	AAR79154 Bovine type IV collagen
7	988	93.9	471 20	AAY44171 Bovine alpha3 type
8	988	93.9	471 21	AAY56783 Partial sequence o
9	986	93.7	471 16	AAR79154

RESULT	1	ALIGNMENTS
ID	AY67942	AY67942 standard; protein; 245 AA.
XX	AY67942;	
AC		03-ARR-2000 (first entry)
XX		
DE		Human type IV collagen alpha 3 chain protein sequence SEQ ID NO:10.
XX		Human; type IV collagen; anti-angiogenic; angiogenesis; cancer;
KW		benign tumour; rheumatoid arthritis; diabetic retinopathy; psoriasis;
KW		ocular angiogenesis disease; Osler-Webber Syndrome; telangiectasia;
KW		myocardial angiogenesis; Plaque neovascularisation; angiofibroma;
KW		atherosclerosis; sclerodema; hypertrophic scar; cat scratch disease;
KW		contraception; obesity.
XX		Homo sapiens.
OS		
XX		
PN		
XX		
PD	23-DEC-1999.	
XX		
PF	17-JUN-1999;	
XX		
PR	99WO-US13737.	
XX		
PR	17-JUN-1998;	
XX		
PR	99US-0089689.	
XX		
PR	25-MAR-1999;	
XX		
PA	(BETH-) BETH ISRAEL DEACONESS MEDICAL CENT.	
XX		
PT		
PI	Kalluri R;	
XX		
DR	WPI; 2000-097708/08.	
XX		
PS	N-PSDB; AA257158.	
XX		
PT	Anti-angiogenic proteins comprising the NCL domain of the alpha 1, 2 or	
3	chain of Type IV collagen used in, e.g. treatment of benign tumors	
and rheumatoid arthritis.		
XX		
XX	Claim 32; Fig 16B; 117pp; English.	
CC	The present sequence represents the human type IV collagen alpha 1, 2 or	
CC	domain of the alpha 1, alpha 2 or alpha 3 chains of type IV collagen or	
CC	a fragment, analogue, derivative or mutant, which has anti-angiogenic	
CC	properties. The anti-angiogenic proteins, multimers and chimeras are	
CC	useful for inhibiting angiogenic activities in mammalian tissue,	
CC	especially for treating diseases chosen from angiogenesis dependent	
CC	cancers, benign tumours, rheumatoid arthritis, diabetic retinopathy,	
CC	psoriasis, ocular angiogenesis, plaque neovascularisation, telangiectasia,	
CC	myocardial angiogenesis, haemophilic angiomas, angiofibroma, wound granulation, intestinal	
CC	adhesions, atherosclerosis, scleroderma, hypertrophic scars, cat scratch	
CC	disease, Helicobacter pylori ulcers, dialysis graft vascular access	
CC	stenosis, contraception and obesity. The compositions can be used to	
CC	inhibit a disease characterised by angiogenic activity, in conjunction	
CC	with radiation therapy, chemotherapy or immunotherapy.	
XX		
SQ	Sequence 245 AA;	
XX		
Query		
Match		
Best Local Similarity	100.0%	score 1052; DB 21; Length 245;
Matches	191;	Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1	QRAHGDLGTLGSLCLQRFTTMLPFLFCNVNDVNFASRNDYSYWLSTPALMPNMAPITGR
Db	55	qrahggdgltgslclqrfttimpfclcnvndvfasrndysylstpalmpnmapitgr 114
QY	61	ALEPY1SRCTCCEGPATAIAVHSQFTDIPPCPHGWISLWKGSFIMTSASSEGTCGALA 120

AC		Query Match		Best Local Similarity		Score		DB 4;		Length		1670;	
				100.0%;		1052;							
				Best Local Similarity		Pred. No.		1.2e-100;					
				Matches		Mismatches		0;		Indels		0;	
R	EMBL; AJ288494;	CAC36101.1;	JOINED.										
R	EMBL; AJ288495;	CAC36101.1;	JOINED.										
R	EMBL; AJ288496;	CAC36101.1;	JOINED.										
R	EMBL; AJ288497;	CAC36101.1;	JOINED.										
R	EMBL; AJ288498;	CAC36101.1;	JOINED.										
R	EMBL; AJ288499;	CAC36101.1;	JOINED.										
R	EMBL; AJ288500;	CAC36101.1;	JOINED.										
R	EMBL; AJ288501;	CAC36101.1;	JOINED.										
R	EMBL; AJ288502;	CAC36101.1;	JOINED.										
R	EMBL; AJ288503;	CAC36101.1;	JOINED.										
R	EMBL; AJ288504;	CAC36101.1;	JOINED.										
R	EMBL; AJ288505;	CAC36101.1;	JOINED.										
R	EMBL; AJ288506;	CAC36101.1;	JOINED.										
R	EMBL; AJ288507;	CAC36101.1;	JOINED.										
R	EMBL; AJ288509;	CAC36101.1;	JOINED.										
R	EMBL; AJ288510;	CAC36101.1;	JOINED.										
R	EMBL; AJ288511;	CAC36101.1;	JOINED.										
R	EMBL; AJ288512;	CAC36101.1;	JOINED.										
R	EMBL; AJ288513;	CAC36101.1;	JOINED.										
R	EMBL; AJ288514;	CAC36101.1;	JOINED.										
R	EMBL; AJ288515;	CAC36101.1;	JOINED.										
R	EMBL; AJ288516;	CAC36101.1;	JOINED.										
R	EMBL; AJ288517;	CAC36101.1;	JOINED.										
R	EMBL; AJ288518;	CAC36101.1;	JOINED.										
R	EMBL; AJ288519;	CAC36101.1;	JOINED.										
R	EMBL; AJ288520;	CAC36101.1;	JOINED.										
R	EMBL; AJ288521;	CAC36101.1;	JOINED.										
R	EMBL; AJ288522;	CAC36101.1;	JOINED.										
R	EMBL; AJ288523;	CAC36101.1;	JOINED.										
R	EMBL; AJ288524;	CAC36101.1;	JOINED.										
R	EMBL; AJ288525;	CAC36101.1;	JOINED.										
R	EMBL; AJ288526;	CAC36101.1;	JOINED.										
R	EMBL; AJ288527;	CAC36101.1;	JOINED.										
R	EMBL; AJ288528;	CAC36101.1;	JOINED.										
R	EMBL; AJ288529;	CAC36101.1;	JOINED.										
R	EMBL; AJ288530;	CAC36101.1;	JOINED.										
R	EMBL; AJ288531;	CAC36101.1;	JOINED.										
R	EMBL; AJ288532;	CAC36101.1;	JOINED.										
R	EMBL; AJ288533;	CAC36101.1;	JOINED.										
R	EMBL; AJ288534;	CAC36101.1;	JOINED.										
R	EMBL; AJ288535;	CAC36101.1;	JOINED.										
R	EMBL; AJ288536;	CAC36101.1;	JOINED.										
R	EMBL; AJ288537;	CAC36101.1;	JOINED.										
Q	SEQUENCE	1670 AA;	161899 MW;	FAT7BEA914CA0A6F6	CRC64;								

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on:

March 6, 2002, 06:54:49 ; Search time 31.16 Seconds
(without alignments)
322.691 Million cell updates/sec

Title: US-09-543-371-10_COPY_1_132
Perfect score: 721
Sequence: 1 GLKGKGDGSPATWTPTRGF.....RALEPYVISRCTVCEGPAIAI 132
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues
To number of hits satisfying chosen parameters: 2

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 90%
Maximum Match 100%
Listing first 1000 summaries

Database : PIR_68:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Score	Match Length	DB ID	Description
1	721	100.0	1670	1 CGHU3B
2	713	98.9	220	2 B49736

ALIGNMENTS

RESULT 1
CGHU3B
collagen alpha 3(IV) chain precursor, long splice form - human
N;Alternate names: Goodpasture antigen; procollagen alpha 3(IV) chain long splice form
C;Keywords: Homo sapiens (man)

C;Date: 28-Oct-1994 #sequence_revision 03-Oct-1995 #text_change 22-Jun-1999
C;Accession: A45763; A43928; A4043; A45971; A39786
R;Mariyama, M.; Leinonen, A.; Mochizuki, T.; Tryggvason, K.; Reeder, S.T.
J;Biol. Chem., 269, 23013-23017, 1994
A;Title: Complete primary structure of the human alpha3(IV) collagen chain. Coexpression
A;Reference number: A54763; MUID:94364994
A;Accession: A54763
A;Molecule type: mRNA
A;Residues: 1-1670 <MAR>
A;Cross-references: GB:x80031; NID:9577563; PID:g577564
A;Experimental source: Kidney
R;Turner, N.; Mason, P.J.; Brown, R.; Fox, M.; Povey, S.; Rees, A.; Pusey, C.D.
A;Title: Molecular cloning of the human Goodpasture antigen demonstrates it to be the al
A;Accession: A43928
A;Molecule type: mRNA

A;Residues: 1331-1524, 'I', 1526-1670 <TUR>
A;Cross-references: GB:M81379
A;Experimental source: kidney
R;Quinones, S.; Bernai, D.; Garcia-Sogo, M.; Elena, S.F.; Saus, J.
J;Biol. Chem., 267, 19780-19784, 1992
A;Title: Exon/intron structure of the human alpha 3(IV) gene encompassing the Goodpasture sequence.
A;Reference number: A44043; MUID:93015826
A;Accession: A44043
A;Molecule type: DNA; mRNA
A;Residues: 1386-1670 <QUIT>
A;Cross-references: GB:M92993; NID:9177895; PIDN:AAA21610.1; PID:9177896
A;Note: sequence extracted from NCBI backbone (NCBIP:113597)
R;Quinones, S.; Bernai, D.; Garcia-Sogo, M.; Elena, S.F.; Saus, J.
J;Biol. Chem., 269, 17358, 1994
A;Reference number: A44738; MUID:94274734
A;Contents: annotation: erratum; correction to intronic sequence in A44043
R;Bernal, D.; Quinones, S.; Saus, J.
J;Biol. Chem., 268, 12090-12094, 1993
A;Title: The human mRNA encoding the Goodpasture antigen is alternatively spliced.
A;Reference number: A45971; MUID:93280184
A;Accession: A45971
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1427-1444 <BER>
A;Note: sequence extracted from NCBI backbone (NCBIP:133363); sequence incorrectly id
R;Morrison, K.E.; Mariyama, M.; Yang-Feng, T.L.; Reeder, S.T.
Am. J. Hum. Genet., 49, 545-554, 1991
A;Title: Sequence and localization of a partial cDNA encoding the human alpha3 chain
A;Reference number: A39786; MUID:91353570
A;Accession: A39786
A;Molecule type: mRNA
A;Residues: 1453-1593, 'A', 1595-1670 <MOR>
A;Cross-references: GB:S5590; NID:933418; PIDN:AB19637.1; PID:9234419
C;Comment: In Goodpasture's syndrome, an autoimmune response develops against an epitope and subsequently O-glycosylated.
C;Keywords:
A;Gene: GDB:COL4A3
A;Cross-references:
A;Introns: 1385/1; 1418/1; 1488/1; 1547/2; 1585/3; 1643/2 #status incomplete *
A;Map Position: 2q36-2q37
A;Product: collagen alpha 3(IV) chain genes are encoded on opposite strands w
C;Complex: This minor type IV collagen is thought to form a heterotrimeric of two alpha
among trimer amino-terminal domains (with disulfide and desmosine cross links), dimer
and associations in the interrupted helical domain (with disulfide and desmosine cross
C;Function:
A;Description: minor structural component of extracellular basement membrane in kidney
C;Superfamily: collagen alpha 1(IV) chain
C;Keywords: alternative splicing; basement membrane; cell binding; coiled coil; extra
F;1-28(Domain: signal sequence #status predicted <SGC>
F;29-1670/Product: collagen alpha 3(IV) chain, long splice form #status predicted <MA
F;29-42(Domain: amino-terminal nonhelical, NHI <NHL>
F;43-1438/Region: interrupted helical
F;791-93/Region: cell attachment (R-G-D) motif
F;1-28(Domain: signal sequence #status predicted <SGC>
F;1154-1156/Region: cell attachment (R-G-D) motif
F;1306-1308/Region: cell attachment (R-G-D) motif
F;1345-1347/Region: cell attachment (R-G-D) motif
F;1432-1434/Region: cell attachment (R-G-D) motif
F;1439-1670/Domain: carboxy-terminal nonhelical, NCI <NCI>
F;1451-1551/Domain: collagen IV carboxyl-terminal repeat <CT1>
F;1561-1665/Domain: collagen IV carboxyl-terminal repeat <CT2>
F;31,33,39,41,125,422,476,479,682,722,809,138/Disulfide bonds: interchain #status pr
F;1460-1548/1493-1551/Disulfide bonds: (or 1460-1551, 1493-1548) #status predicted
F;1505-1511,1616-1622/Disulfide bonds: #status predicted
F;1570-1662,1604-1665/Disulfide bonds: (or 1570-1665, 1604-1662) #status predicted
Query Match 100.0%; Score 721; DB 1; Length 1670;
Best Local Similarity 100.0%; Pred. No. 1 8e-65;
Matches 132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GLGKRGDGSATWTRGFVFRHSSATAIPSCPEGVPLYSGFSFLFVQGNQRAHQGD 60
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 Db 1427 GLRKRGDGSATWTRGFVFRHSSATAIPSCPEGVPLYSGFSFLFVQGNQRAHQGD 1486

Qy 61 LGTGLSCLQRFTMPFLFCNVNDVCNFASRNDYWSLTPALMMAPITGAGLEPYIS 120
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 Db 1487 LGTGLSCLQRFTMPFLFCNVNDVCNFASRNDYWSLTPALMMAPITGAGLEPYIS 1546

Qy 121 RCTVCEGPAIAI 132
 ||||||| |||||||

Db 1547 RCTVCEGPAIAI 1558

RESULT 2

B49736 collagen alpha 3(IV) chain, medium splice form - human (fragment)
 Contains: collagen alpha 3(IV) chain, splice form GP-V

C;Species: Homo sapiens (man)
 Date: 03-May-1994 #sequence_revision 12-Nov-1999 #text_change 17-Mar-2000

Accession: B49736; D49736; S69111

R;Feng, L.; Xia, Y.; Wilson, C.B.

J. Biol. Chem. 269: 2342-2348, 1994

A;Title: Alternative Splicing of the N1 domain of the human alpha3(IV) collagen gene.

A;Reference number: A49736; MUID:94124597

A;Status: nucleic acid sequence not shown

A;Molecule type: mRNA

A;Residues: 169-220 <FPN1>

A;Accession: D49736
 A;Status: nucleic acid sequence not shown; translation not shown

A;Molecule type: mRNA

A;Residues: 22-220 <FPN2>

A;Cross-references: GB02519; NID:9409106; PIDN:AAA18942.1; PID:9409107

A;Note: this is the conceptual translation of the nucleic acid submitted to GenBank

R;Pendes, J.R.; Bernal, D.; Revart, F.; Johansson, C.; Fresquet, V.J.; Cervera, J.; Wie

Eur. J. Biochem. 229: 754-760, 1995

A;Title: Characterization and expression of multiple alternatively spliced transcripts c

A;utotantigen and one of its alternative forms.

A;Reference number: S69111; MUID:95278230

A;Accession: S69111
 A;Molecule type: mRNA

A;Residues: 1-45, 169-204, 'L', 206-220 <PEN>

C;Comment: For the complete sequence of the long splice form, see PIR:CGH03B.

C;Genetics: GDB:COL4A3

A;Cross-references: GDB:128351; OMIM:120070

A;P position: 2a36-2a37
 A;pefamily: collagen alpha 1(IV) chain

C;Keywords: alternative splicing; basement membrane; cell binding; coiled coil; extracellular

F;1-220;Product: collagen alpha 3(IV) chain, medium splice form (fragment) #status pred;
 F;1-45, 169-220;Product: collagen alpha 3(IV) chain, splice from GP-V (fragment) #status

F;22-220;Domain: carboxyl-terminal nonhelical, N1 <ncl>
 F;34-134;Domain: collagen IV carboxyl-terminal repeat <ct1>

Query Match 98.9%; Score 713; DB 2; Length 220;
 Best Local Similarity 99.2%; Pred. No. 1.4e-65; Mismatches 131; Conservative 0; Indels 0; Gaps 0;
 Matches 131; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GLKKGKRGDGSATWTRGFVFRHSSATAIPSCPEGVPLYSGFSFLFVQGNQRAHQGD 60
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 Db 10 GLKKGKRGDGSATWTRGFVFRHSSATAIPSCPEGVPLYSGFSFLFVQGNQRAHQGD 69

Qy 61 LGTGLSCLQRFTMPFLFCNVNDVCNFASRNDYWSLTPALMMAPITGAGLEPYIS 120
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 Db 70 LGTGLSCLQRFTMPFLFCNVNDVCNFASRNDYWSLTPALMMAPITGAGLEPYIS 129

Qy 121 RCTVCEGPAIAI 132
 ||||||| |||||||

Db 130 RCTVCEGPAIAI 141

Search completed: March 6, 2002, 06:54:49
 Job time: 40 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on:

March 6, 2002, 07:05:08 ; Search time 18.75 Seconds

(without alignments)
258.120 Million cell updates/sec

Title: US-09-543-371-10_COPY_1_132

Perfect score: 721

Sequence: 1 GLKQKRGDSGSPATWTRGF.....RALEPYVISRCTIVCEGPAIAI 132

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 10059 seqs, 36664827 residues

T number of hits satisfying chosen parameters: 1

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 90%
Maximum Match 100%

Listing first 1000 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Match Length	DB ID	Description
1	721	100.0	1670	CA34_HUMAN

ALIGNMENTS

RESULT	1	CA34_HUMAN	STANDARD;	PRT:	1670 AA.
CA34_HUMAN	001955;	D	01-OCT-1996 (Rel. 34, Created)		
	DT	15-JUL-1999 (Rel. 38, Last sequence update)			
	DT	20-AUG-2001 (Rel. 40, Last annotation update)			
DE	COLLAGEN ALPHA 3(IV) CHAIN PRECURSOR (GOODPASTURE ANTIGEN).				
GN	COL4A3.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniate; Vertebrata; Buteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Kidney;				
RX	Turner N., Mason P.J., Brown R., Fox M., Povey S., Rees A.,				
RA	Pussy C.D.;				
RT	"Molecular cloning of the human Goodpasture antigen demonstrates its 3 chain of type IV collagen.";				
RL	J. Clin. Invest. 89:592-601(1992).				
RN	[16]				
RP	SEQUENCE OF 1644-1670 FROM N.A.				
RC	TISSUE=Kidney;				
RX	Ding J.;				
RN	[17]				
RP	Submitted (JAN-1993) to the EMBL/GenBank/DBJ databases.				
RN	[18]				
RP	SEQUENCE OF 1-239 FROM N.A.				
RX	Feng L., Xia Y., Wilson C.B.;				
RT	"Alternative splicing of the NC1 domain of the human alpha 3(IV) collagen gene: differential expression of mRNA transcripts that predict three protein variants with distinct carboxyl regions.";				
RT	J. Biol. Chem. 269:2342-2348(1994).				
RX	MEDLINE=94124597; PubMed=8294492;				
RX	Feng L., Xia Y., Wilson C.B.;				
RT	"The human mRNA encoding the Goodpasture antigen is alternatively spliced.";				
RT	J. Biol. Chem. 268:12090-12094(1993).				
RN	[19]				
RP	ALTERNATIVE SPLICING.				
RX	MEDLINE=93280184; PubMed=8505332;				
RA	Bernal D., Quinones S., Saus J.;				
RT	"The human mRNA encoding the Goodpasture antigen is alternatively spliced.";				
RT	J. Biol. Chem. 268:12090-12094(1993).				
RN	[10]				
RP	VARIANT PRO-1474.				
RX	MEDLINE=95078827; PubMed=7987301;				
RA	Leimink H.H., Mochizuki T., van den Heuvel L.P.W.J., Schroeder C.H.,				
RA	Barrientos A., Monnens L.A.H., van Oost B.A., Brunner H.G.,				
RA	Reeders S.T., Smeets H.J.M.;				
RT	"Mutations in the type IV collagen alpha 3 (COL4A3) gene in autosomal recessive Alport syndrome.";				
RT	J. Hum. Mol. Genet. 3:1269-1273(1994).				
RN	[11]				
RP	VARIANTS AS, AND VARIANTS.				
RX	MEDLINE=21064696; PubMed=11134255;				
RA	Heijet L., Artondel C., Forrestier L., Cohen-Solal L., Mollet G.,				
RA	Gutierrez B., Stavrou C., Guiller M.C., Antignac C.;				
RT	"Structure of the human type IV collagen gene COL4A3 and mutations in autosomal Alport syndrome.";				
RT	J. Am. Soc. Nephrol. 12:97-105 (2001).				
RL	Leinonen A.;				
CC	Glomerular Basement Membranes (GBM), forming a 'chicken-wire' meshwork together with laminins, proteoglycans and entactin/				
CC	[3] Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.				

CC NITROGEN.
 CC - I - SUBUNIT: THERE ARE SIX TYPE IV COLLAGEN ISOFORMS, ALPHA 1(IV)-
 CC ALPHA 6(IV), EACH OF WHICH CAN FORM A TRIPLE HELIX STRUCTURE
 CC WITH 2 OTHER CHAINS TO GENERATE TYPE IV COLLAGEN NETWORK.
 CC - I - SUBCELLULAR LOCATION: CELL SURFACE (POTENTIAL).
 CC - I - ALTERNATIVE PRODUCTS: AT LEAST 3 ISOFORMS: 1 (SHOWN HERE), 2/V AND
 CC 3/L5, ARE PRODUCED BY ALTERNATIVE SPlicing. THEY DIFFER IN THEIR
 CC C-TERMINAL N-TERMINAL N-TERMINAL N-TERMINAL N-TERMINAL N-TERMINAL
 CC - I - TISSUE SPECIFICITY: ALPHA 3 AND ALPHA 4 TYPE IV COLLAGENS ARE
 CC COLOCALIZED AND PRESENT ONLY IN BASEMENT MEMBRANES OF KIDNEY, EYE,
 CC COCHLEA, LUNG AND BRAIN.
 CC - I - DOMAIN: ALPHA CHAINS OF TYPE IV COLLAGEN HAVE A NONCOLLAGENOUS
 DOMAIN (NC1) AT THEIR C-TERMINUS, FREQUENT INTERRUPTIONS OF THE
 G-X-Y REPEATS IN THE LONG CENTRAL TRIPLE-HELICAL DOMAIN (WHICH MAY
 CAUSE FLEXIBILITY IN THE TRIPLE HELIX), AND A SHORT N-TERMINAL
 TRIPLE-HELICAL 7S DOMAIN.
 CC - I - PTM: PROLINES AT THE THIRD POSITION OF THE TRIPETIDE REPEATING
 UNIT (G-X-Y) ARE HYDROXYLATED IN SOME OR ALL OF THE CHAINS.
 CC - I - PTM: THE ALTERNATIVE SPLICED FORM V CONTAINS AN ADDITIONAL
 N-LINKED GLYCOSYLATION SITE.
 CC - I - PTM: TYPE IV COLLAGENS CONTAIN NUMEROUS CYSTEINE RESIDUES WHICH
 ARE INVOLVED IN INTRAMOLECULAR DISULFIDE BONDING. 12 OF
 THESE, LOCATED IN THE NC1 DOMAIN, ARE CONSERVED IN ALL KNOWN TYPE
 IV COLLAGENS.
 CC - I - DISEASE: ANTIBODIES AGAINST THE NC1 DOMAIN OF ALPHA3(IV) MEDIATE
 THE HUMAN AUTOIMMUNE DISEASE, GOODPASTURE SYNDROME, WHICH IS
 CHARACTERIZED BY HEMATURIA AND PULMONARY HEMORRHAGE.
 CC - I - DISEASE: DEFECTS IN COLA3 ARE ASSOCIATED WITH THE TYPE I
 AUTOSOMAL RECESSIVE FORM OF ALPORT SYNDROME, AN HEREDITARY
 GLOMERULONEPHRATIC CHARACTERIZED BY PROGRESSIVE RENAL FAILURE,
 HEMATURIA AND DEAFNESS. THE RECESSIVE FORM OCCURS EQUALLY BETWEEN
 MALES AND FEMALES.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the BMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC
 CC
 DR EMBL: XH0031; CAAX6335; 1; -
 DR EMBL: M22993; AA31610; 1; -
 DR EMBL: S55790; AAB19637; 1; -
 DR EMBL: M31379; AA31556; 1; -
 DR EMBL: LA8650; AA512044; 1; -
 DR EMBL: U02519; AAAB18942; 1; -
 DR EMBL: U02520; AAAB18943; 1; -
 DR MIM: 120070; -
 DR MIM: 203780; -
 DR MIM: 231450; -
 DR InterPro: IPR001442; C4
 DR InterPro: IPR000087; Collagen.
 DR Pfam: PF01413; C4; 2.
 DR SMART: SM00111; C4; 2.
 DR Pfam: PF01391; Collagen; 21.
 DR ProDom: PD005923; C4; 2.
 DR SMART: SM00111; C4; 2.
 KW Extracellular matrix; Connective tissue; Repeat; Hydroxylation;
 KW Glycoprotein; Basement membrane; Collagen; Signal; Cell adhesion;
 KW Alternative splicing; Polymorphism; Phosphorylation; Disease mutation;
 KW Airport syndrome.
 FT SIGNAL 1 28 POTENTIAL.
 FT CHAIN 29 1670 COLLAGEN ALPHA 3 (IV) CHAIN.
 FT DOMAIN 29 42 7S DOMAIN.
 FT DOMAIN 43 1438 TRIPLE-HELICAL REGION.
 FT DOMAIN 1439 1670 NONHELIICAL REGION (NC1) (GOODPASTURE
 FT ANTIGEN) (BY SIMILARITY).
 FT DOMAIN 1427 1444 EPITOPE RECOGNIZED BY GOODPASTURE
 FT ANTIBODIES.
 FT SITE 1426 1427 CLEAVAGE (BY COLLAGENASE)
 (BY SIMILARITY).
 FT SITE 791 793 CELL ATTACHMENT SITE (POTENTIAL).

FT SITE 996 998 CELL ATTACHMENT SITE (POTENTIAL).
 FT SITE 1154 1156 CELL ATTACHMENT SITE (POTENTIAL).
 FT SITE 1306 1308 CELL ATTACHMENT SITE (POTENTIAL).
 FT SITE 1345 1347 CELL ATTACHMENT SITE (POTENTIAL).
 FT SITE 1432 1434 CELL ATTACHMENT SITE (POTENTIAL).
 FT CARBOHYD 253 253 N-LINKED (GLICNAc, . .) (POTENTIAL).
 FT MOD-RES 1435 1435 PHOSPHORYLATION (BY SIMILARITY).
 FT MOD-RES 1437 1437 PHOSPHORYLATION (BY SIMILARITY).
 FT DISULFID 1460 1551 OR 1548 (BY SIMILARITY).
 FT DISULFID 1493 1548 OR 1551 (BY SIMILARITY).
 FT DISULFID 1505 1511 OR 1662 (BY SIMILARITY).
 FT DISULFID 1604 1662 OR 1665 (BY SIMILARITY).
 FT DISULFID 1616 1622 BY SIMILARITY.
 FT VARSSPLIC 1586 1670 PTM SEQUENCE: GTCGSCLCQRTTMPLFCNVNDCNFAASRNDSYWLSPVAKGELEKITSRCWCMK
 KKH -> KAYSINCSEGWRKNSLSSVHEEKTLLKAKTA
 ELVEFFKNUKVKMTHAVI (IN ISOFORM 2).
 GTGCGSCLCQRTTMPLFCNVNDCNFAASRNDSYWLSPVAKGELEKITSRCWCMK
 MPMMNAPITGRALEPKISRCWCMKPTVCEGGATIAHVSQDTPP
 CPGKWTISLMKGFSITMFNSAGSGCTGQALASPCLEFERA
 SPELECHGGTCNYSNSYNSFWLNLNPERMRKPIPTSVK
 AGEELERK11SRCQVNMKKRH -> DALFKVLRSP (IN
 ISOFORM 3).
 G -> R.
 /FT id=VAR_011202.
 G -> E.
 /FT id=VAR_011203.
 G -> E (IN AS).
 FT VARIANT 297 297 Query Match 100.0% Score 721; DB 1; Length 1670;
 Best Local Similarity 100.0%; Pred. No. 7.9e-68; Indels 0; Gaps 0;
 Matches 132; Conservative 0; Mismatches 0; Gaps 0;
 Db 1427 GLKGKRGDSGSPATWITRQFVFRHSQTAIPSCPEGVWPLYSGFSLFLVQGNQRAGQD 60
 Qy 61 LGTIGSCLQRTTMPLFCNVNDCNFAASRNDSYWLSPVAKGELEKITSRCWCMK 1486
 Db 1487 LGTIGSCLQRTTMPLFCNVNDCNFAASRNDSYWLSPVAKGELEKITSRCWCMK 120
 Qy 121 RCTVCEGPVIAI 132
 Db 1547 RCTVCEGPVIAI 1558

Search completed: March 6, 2002, 07:05:09
 Job time: 585 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 6, 2002, 07:04:44 ; Search time 50.56 Seconds
(without alignments)

US-09-543-371-10_COPY_1_132
381.882 Million cell updates/sec

Title: perfect score:
Sequence: 1 GLKGKRGDGSQPATWTRGF.....RALEPYISRCTVCEGPAIAI 132
Scoring table: BLOSUM62
Gap0 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

To number of hits satisfying chosen parameters: 2

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 90%
Maximum Match 100%

Listing first 1000 summaries

Database : SPREMBL_17:*

- 1: sp_archea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mic:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	Query	Score	Match	Length	DB	ID	Description
1	9NYC4	100.0	245	4	9NYC4		09nyc4 homo sapien
2	9NYC4	100.0	1670	4	Q9bqt2		Q9bqt2 homo sapien

ALIGNMENTS

RESULT	1
Q9NYC4	PRELIMINARY; PRT; 245 AA.
ID	09NYC4;
AC	09NYC4;
DT	01-OCT-2000 (TREMBLrel. 15, Created) 01-OCT-2000 (TREMBLrel. 15, Last sequence update) 01-OCT-2000 (TREMBLrel. 17, Last annotation update)
DR	01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE	TUMSTATIN (FRAGMENT).
OS	Homo sapiens (Human)
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX	NCBI_TaxID=9606;

RN	[1]
SEQUENCE FROM N.A.	
RP	
RA	
RA	Maeshima Y., Colorado P.C., Torre A., Holthaus K.A., Grunkemeyer J.A.,
RA	Ericken M.D., Hopfer H., Xiao Y., Stillman I.E., Kalluri R.,
RT	"Distinct anti-tumor properties of a type IV collagen domain derived from basement membrane.";
RL	J. Biol. Chem. 0:0-0(2000);
EMBL	AF258351; AAF2632.1; -.
DR	Interrr; IPR001442; C4.
DR	Interpro; IPR000504; RRM.
DR	pFam; PF01413; C4; 2.
DR	ProDom; PDO0392; C4; 2.
DR	SMART; SM00111; C4; 2.
PROSITE; PS00030; RRM_RNP_1; UNKNOWN_1.	
FT	NON_TER 1
SQ	SEQUENCE 245 AA; 26952 MW; 1EE5028354D9A57D CRC64;
Query	GLKGKRGDGSQPATWTRGF
Match	100.0%; score 721; DB 4; length 245;
Best Local Similarity	100.0%; Pred. No. 2.4e-72;
Matches	132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1
Db	2 GLKGKRGDGSQPATWTRGF
QY	61 LGTLGSCLQRRTMPFLFCNNNDVCINFASRNDRYSYMLSTPAMPMMAPITGRALEPYIS
Db	62 LGTLGSCLQRRTMPFLFCNNNDVCINFASRNDRYSYMLSTPAMPMMAPITGRALEPYIS
QY	121 RCTVCEGPAIAI 132
Db	122 RCTVCEGPAIAI 133
RESULT	2
O9BQT2	O9BQT2
ID	O9BQT2
AC	O9BQT2;
DT	01-JUL-2001 (TREMBLrel. 17, Created)
DT	01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT	01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE	ALPHA3 TYPE IV COLLAGEN.
GN	COL4A3.
OS	Homo sapiens (Human)
OC	Bukarota; Metacoda; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX	NCBI_TaxID=9606;
RN	[1]
SEQUENCE FROM N.A.	
RP	SEQUENCE FROM N.A.
RX	MEDLINE=21064695; PubMed=11134255;
RA	Heidet L., Arrendel C., Forestier L., Cohen-Solal L., Mollet G.,
RA	Gutierrez B., Savrou C., Gubler M.C., Antignac C.,
RT	"Structure of the human type IV collagen gene COL4A3 and mutations in
RT	autosomal Alport syndrome";
RT	J. Am. Soc. Nephrol. 12:97-106 (2001).
RL	EMBL; AJ288487; CAC36101.1; -.
DR	EMBL; AJ288488; CAC36101.1; JOINED.
DR	EMBL; AJ288489; CAC36101.1; JOINED.
DR	EMBL; AJ288490; CAC36101.1; JOINED.
DR	EMBL; AJ288491; CAC36101.1; JOINED.
DR	EMBL; AJ288492; CAC36101.1; JOINED.
DR	EMBL; AJ288493; CAC36101.1; JOINED.
DR	EMBL; AJ288494; CAC36101.1; JOINED.
DR	EMBL; AJ288495; CAC36101.1; JOINED.
DR	EMBL; AJ288496; CAC36101.1; JOINED.
DR	EMBL; AJ288497; CAC36101.1; JOINED.
DR	EMBL; AJ288498; CAC36101.1; JOINED.
DR	EMBL; AJ288499; CAC36101.1; JOINED.
DR	EMBL; AJ288501; CAC36101.1; JOINED.
DR	EMBL; AJ288502; CAC36101.1; JOINED.
DR	EMBL; AJ288503; CAC36101.1; JOINED.

DR EMBL; AJ288504; CAC36101.1; JOINED.
 DR EMBL; AJ288505; CAC36101.1; JOINED.
 DR EMBL; AJ288506; CAC36101.1; JOINED.
 DR EMBL; AJ288507; CAC36101.1; JOINED.
 DR EMBL; AJ288508; CAC36101.1; JOINED.
 DR EMBL; AJ288509; CAC36101.1; JOINED.
 DR EMBL; AJ288510; CAC36101.1; JOINED.
 DR EMBL; AJ288511; CAC36101.1; JOINED.
 DR EMBL; AJ288512; CAC36101.1; JOINED.
 DR EMBL; AJ288513; CAC36101.1; JOINED.
 DR EMBL; AJ288514; CAC36101.1; JOINED.
 DR EMBL; AJ288515; CAC36101.1; JOINED.
 DR EMBL; AJ288516; CAC36101.1; JOINED.
 DR EMBL; AJ288517; CAC36101.1; JOINED.
 DR EMBL; AJ288518; CAC36101.1; JOINED.
 DR EMBL; AJ288519; CAC36101.1; JOINED.
 DR EMBL; AJ288520; CAC36101.1; JOINED.
 DR EMBL; AJ288521; CAC36101.1; JOINED.
 DR EMBL; AJ288522; CAC36101.1; JOINED.
 DR EMBL; AJ288523; CAC36101.1; JOINED.
 DR EMBL; AJ288524; CAC36101.1; JOINED.
 DR EMBL; AJ288525; CAC36101.1; JOINED.
 DR EMBL; AJ288526; CAC36101.1; JOINED.
 DR EMBL; AJ288527; CAC36101.1; JOINED.
 DR EMBL; AJ288528; CAC36101.1; JOINED.
 DR EMBL; AJ288529; CAC36101.1; JOINED.
 DR EMBL; AJ288530; CAC36101.1; JOINED.
 DR EMBL; AJ288531; CAC36101.1; JOINED.
 DR EMBL; AJ288532; CAC36101.1; JOINED.
 DR EMBL; AJ288533; CAC36101.1; JOINED.
 DR EMBL; AJ288534; CAC36101.1; JOINED.
 DR EMBL; AJ288535; CAC36101.1; JOINED.
 DR EMBL; AJ288536; CAC36101.1; JOINED.
 DR EMBL; AJ288537; CAC36101.1; JOINED.
 DR EMBL; AJ288538; CAC36101.1; JOINED.
 KW Collagen.

SQ SEQUENCE 1670 AA; 161899 MW; FA7BB4914CA0A6F6 CRC64;

Query Match 100.0%; Score 721; DB 4; Length 1670;
 Best Local Similarity 100.0%; Pred. No. 1 9e-71;
 Matches 132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GLKGKRGDSGSPATWTRGFVTRHSSQTAIPSCPEGVPLYSGFSFLFVGQNRAHGQD 60
 Db 1427 GLKGKRGDSGSPATWTRGFVTRHSSQTAIPSCPEGVPLYSGFSFLFVGQNRAHGQD 1486
 61 LGTIGSCLQRFITMPFELCNVNDVCNFSRNDSYWLSTPALPMNMAPITGRALEPYIS 120
 Db 1487 LGTIGSCLQRFITMPFELCNVNDVCNFSRNDSYWLSTPALPMNMAPITGRALEPYIS 1546
 QY 121 RCTVCEGPAl 132
 Db 1547 RCTVCEGPAl 1558

Gencore version 4.5
 Copyright (c) 1993 - 2000 compugen Ltd.

OM protein - protein search, using sw model

Run on: March 6, 2002, 06:56:24 ; Search time 54.67 Seconds
 (without alignments)
 178.849 Million cell updates/sec

Title: US-09-543-371-10_COPY_1_132

Perfect score: 721
 Sequence: 1 GLKGRKGDSGSPATWTRGFV... RALEPYIISRCTVCEGPAIAI 132

Scoring table: BLOSUM62
 Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 4

Minimum DB seq length: 0
 Maximum DB seq length: 200000000

Post-processing: Minimum Match 90%
 Maximum Match 100%
 Listing first 1000 summaries

Database : A_Geneseq_1101:*

1: /SIDS8/gcadata/geneseq/geneseq/AA1980.DAT: *
 2: /SIDS8/gcadata/geneseq/geneseq/AA1981.DAT: *
 3: /SIDS8/gcadata/geneseq/geneseq/AA1982.DAT: *
 4: /SIDS8/gcadata/geneseq/geneseq/AA1983.DAT: *
 5: /SIDS8/gcadata/geneseq/geneseq/AA1984.DAT: *
 6: /SIDS8/gcadata/geneseq/geneseq/AA1985.DAT: *
 7: /SIDS8/gcadata/geneseq/geneseq/AA1986.DAT: *
 8: /SIDS8/gcadata/geneseq/geneseq/AA1987.DAT: *
 9: /SIDS8/gcadata/geneseq/geneseq/AA1988.DAT: *
 10: /SIDS8/gcadata/geneseq/geneseq/AA1989.DAT: *
 11: /SIDS8/gcadata/geneseq/geneseq/AA1990.DAT: *
 12: /SIDS8/gcadata/geneseq/geneseq/AA1991.DAT: *
 13: /SIDS8/gcadata/geneseq/geneseq/AA1992.DAT: *
 14: /SIDS8/gcadata/geneseq/geneseq/AA1993.DAT: *
 15: /SIDS8/gcadata/geneseq/geneseq/AA1994.DAT: *
 16: /SIDS8/gcadata/geneseq/geneseq/AA1995.DAT: *
 17: /SIDS8/gcadata/geneseq/geneseq/AA1996.DAT: *
 18: /SIDS8/gcadata/geneseq/geneseq/AA1997.DAT: *
 19: /SIDS8/gcadata/geneseq/geneseq/AA1998.DAT: *
 20: /SIDS8/gcadata/geneseq/geneseq/AA1999.DAT: *
 21: /SIDS8/gcadata/geneseq/geneseq/AA2000.DAT: *
 22: /SIDS8/gcadata/geneseq/geneseq/AA2001.DAT: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	721	100.0	211	Human Goodpasture
2	721	100.0	245	Human type IV collagen
3	700	97.1	268	Type IV collagen N
4	700	97.1	21	Human alpha3(IV)NC

ALIGNMENTS

RESULT 1
 AAY95918

Query	Match	Score	Length	DB ID	Description
GLKGRKGDSGSPATWTRGFV... RALEPYIISRCTVCEGPAIAI	100.0%	721	21	211	Human Goodpasture
GLKGRKGDSGSPATWTRGFV... RALEPYIISRCTVCEGPAIAI	100.0%	121	132	RCVCEGPAIAI 132	Human type IV collagen
GLKGRKGDSGSPATWTRGFV... RALEPYIISRCTVCEGPAIAI	100.0%	121	132	rcvcegpaiai 132	Type IV collagen N
GLTGSCLQRFTMPFLFCNVNNDVCNFASRDNYSWLMSTPAPMPMNAAPTGRALEYIS	100.0%	61	120	lgfLgscIqrftmpflfcnvndvcnfasnrdyswlmstpalmpmnapitgraleypis 120	Human alpha3(IV)NC
lgfLgscIqrftmpflfcnvndvcnfasnrdyswlmstpalmpmnapitgraleypis	100.0%	61	120	lgfLgscIqrftmpflfcnvndvcnfasnrdyswlmstpalmpmnapitgraleypis 120	
AY67942	100.0%	121	121	AY67942 standard; Protein; 245 AA.	ID AAY95918;
AY67942	100.0%	121	121	AY67942 standard; Protein; 245 AA.	ID AAY95918;

CC facilitate protein secretion, and a mature protein comprising an affinity tag (facilitates purification and identification of the material) and the alpha-1 chain monomer. The invention provides methods and kits for inhibiting angiogenesis, tumour growth and metastasis, and endothelial cell interaction with the extracellular matrix, each method comprising contacting the tumour or extracellular tissue with 1 or more isolated type IV collagen NCI alpha chain monomer(s) selected from the group consisting of alpha-1, alpha-2, alpha-3 and alpha-6 NCI chain monomers (see AAY31991-96). The monomers can be produced via recombinant protein expression. The polypeptides and polypeptides are used to treat an angiogenesis-mediated disorder or condition, especially selected from solid and blood-borne tumours, diabetic retinopathy, rheumatoid arthritis, retinal neovascularization, choroidal neovascularization, macular degeneration, corneal graft rejection, neovascular glaucoma, retrobulbar fibroplasia, epidemic keratoconjunctivitis, vitamin A deficiency, contact lens overwear, atopic keratitis, superior limbic keratitis, pterygium keratitis sicca, sogners, acne rosacea, phylectenulosis, syphilis, mycobacteria infections, lipid degeneration, chemical burns, bacterial ulcers, fungal ulcers, herpes simplex infections, herpes zoster infections, protozoan infections, Kaposi's sarcoma, Moerens ulcer, Terrien's marginal degeneration, marginal keratolysis, trauma, systemic lupus, polyarteritis, Wegener's sarcoidosis, scleritis, Steven's Johnson disease, radial keratotomy, sickle cell anaemia, sarcoid, pseudoxanthoma elasticum, Paget's disease, vein occlusion, artery occlusion, carotid obstructive disease, chronic uveitis, chronic vitritis, Lyme's disease, Bechets disease, myopia, optic pits, Stargarts disease, pars planitis, chronic retinal detachment, hyperviscosity syndrome, toxoplasmosis, post-laser complications, abnormal proliferation of fibrovascular tissue, haemangiomas, Osler-Weber-Rendu, ocular neovascular disease, osteoarthritis, chronic inflammation, Crohn's disease, ulcerative colitis, psoriasis, atherosclerosis, and pemphigoid (all claimed).

XX Sequence 268 AA:

Query Match 97.1%; Score 700; DB 20; Length 268;
Best Local Similarity 100.0%; Pred. No. 4.8e-70;
Matches 128; Conservative 0; Mismatches 0; Indels 0; Caps 0;

QY 5 KRGDSSGSPATWTRGVFTRHSQTTAIPSCPEGTPLVLYSGESFLVFQGNQRAGHDGTL 64
29 krgdsgspatwtgrgvfrhsqttalpscpegtplvlysgifslfvqgnqrhggdgtl 88

QY 65 GSCLQRTTMAPLFCNVNDVCNFAASRDYWSWLSTPALMPNMAPTGRALEPYISRCTV 124
D 89 gsclqrftttmpfifcnvdvcfasndyswvlstpalmppnmapitgralepyisrcy 148

QY 125 CRGPAAI 132
Db 149 cegpaiai 156

RESULT 4
ID AAY97555 standard; Protein: 268 AA.
AC AAY97555;
XX DT 12-FEB-2001 (first entry)
DE Human alpha3(IV)NC1 protein sequence.

KW chronic inflammation; psoriasis; therapy; alpha3(IV)NC1.
XX OS Homo sapiens.
XX PN WO20059532-A1.
XX PD 12-OCT-2000.
XX PF 31-MAR-2000; 2000WO-US08678.
XX PR 01-APR-1999; 99US-0127391.
XX PA (BIOS-) BIOSTRACTUM INC.
XX PI Brooks P, Hudson B;
DR WPI; 2000-664962/64.
N-PSDB; AAY31993.

SQ Use of antagonists of specific integrin receptors for inhibiting angiogenesis, tumour growth or metastases, or endothelial cell interactions with the extracellular matrix Disclosure: Fig 17c; 78pp; English.

CC This sequence is a human type IV collagen alpha chain monomer, designated alpha3(IV)NC1. The invention relates to a method for inhibiting angiogenesis, tumour growth or metastases, or endothelial cell interactions with the extracellular matrix, comprising contacting the cells or tissue with a polypeptide composition containing antagonists of specific integrin receptors. The methods and the antagonists are useful for inhibiting angiogenesis, tumour growth or metastases, or endothelial cell interaction with the extracellular matrix. The antagonists are also useful for treating diseases and conditions with accompanying undesired angiogenesis, e.g. solid and blood-borne tumours (e.g. melanomas, carcinomas, sarcomas, rhabdomyosarcoma, retinoblastoma, Ewing sarcoma, neuroblastoma, osteosarcoma or leukemia). These are also applicable to treating non-tumorigenic diseases and conditions with accompanying undesired angiogenesis, e.g. diabetic retinopathy, rheumatoid arthritis, retinal neovascularisation, choroidal neovascularisation, muscular degeneration, corneal graft rejection, vitritis, chemical burns, Kaposi's sarcoma, keratitis, Mycobacteria infections, sickle cell anaemia, sarcoid, carotid obstructive disease, post-laser complications, chronic inflammation or psoriasis.

XX Sequence 268 AA:

Query Match 97.1%; Score 700; DB 21; Length 268;
Best Local Similarity 100.0%; Pred. No. 4.8e-70;
Matches 128; Conservative 0; Mismatches 0; Indels 0; Caps 0;

QY 5 KRGDSSGSPATWTRGVFTRHSQTTAIPSCPEGTPLVLYSGESFLVFQGNQRAGHDGTL 64
29 krgdsgspatwtgrgvfrhsqttalpscpegtplvlysgifslfvqgnqrhggdgtl 88

QY 65 GSCLQRTTMAPLFCNVNDVCNFAASRDYWSWLSTPALMPNMAPTGRALEPYISRCTV 124
D 89 gsclqrftttmpfifcnvdvcfasndyswvlstpalmppnmapitgralepyisrcy 148

QY 125 CRGPAAI 132
Db 149 cegpaiai 156

Search completed: March 6, 2002, 06:56:24
Job time: 135 sec

KW Type IV collagen alpha chain monomer; human; inhibitor; angiogenesis; tumour growth; integrin receptor; carcinoma; sarcoma; rhabdomyosarcoma; retinoblastoma; Ewing sarcoma; neuroblastoma; osteosarcoma; leukaemia; diabetic retinopathy; rheumatoid arthritis; neovascularisation; muscular degeneration; corneal graft rejection; vitamin A deficiency; atopic keratitis; Mycobacteria infection; chemical burn; sarcoid; Kaposi's sarcoma; sickle cell anaemia; carotid obstructive disease;

THIS PAGE BLANK (USPTO)

Copyright (c) 1993 - 2000 Compugen Ltd.
GenCore version 4.5

OM protein - protein search, using sw model

Run on:

March 6, 2002, 06:55:22 ; Search time 26.9 Seconds
(without alignments)
110.425 Million cell updates/sec

Title: US-09-543-371-10_COPY_1_132

Perfect score: 721

Sequence: 1 GLKGKRGDGSQSPATWTTTGF.....RALEPYVISRCTVCEGPAIAI 132

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 0

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 90%

Maximum Match 100%

Listing first 1000 summaries

Database : Issued_Patents_AA:*

```

1: /cgn2_6/ptodata/2/iaa/5A_COMBO.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMBO.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMBO.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMBO.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMBO.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description

No matches found

Search completed: March 6, 2002, 06:55:22
Job time: 73 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 6, 2002, 06:54:49 ; Search time 31.16 Seconds

(without alignments)
156.456 Million cell updates/sec

Title: US-09-543-371-10_COPY_181_244

Perfect score: 353

Sequence: 1 EFRASPFLECHGRGTCNVYS..... KAGELEKLISRQCVCKKKRH 64

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 3

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 90%
Maximum Match 100%

Listing first 1000 summaries

Database : PIR:68;*

1: pir1;*

2: pir2;*

3: pir3;*

4: pir4;*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match Length	DB ID	Description
1	353	100.0	1670	1 CGHU3B	collagen alpha 3(I)
2	345	97.7	246	2 I48302	collagen alpha 3(I)
3	323	91.5	471	2 A39024	collagen alpha 3(I)

ALIGNMENTS

RESULT	1	CGHU3B	N	Alternate names: Goodpasture antigen; procollagen alpha 3(IV) chain long splice form - human collagen alpha 3(IV) chain precursor, long splice form
C;Species: Homo sapiens (man)			C;Species: Homo sapiens (man)	
C;Date: 28-Oct-1994 #sequence_revision:03-Oct-1995 #text_change:22-Jun-1999			C;Complex: This minor type IV collagen is thought to form a heterotrimer of two alpha monomer amino-terminal domains (with disulfide and desmosine cross-links), dimeric associations in the interrupted helical domain (with disulfide and desmosine cross-	
C;Accession: A54763; A43928; A40434; A45971; A39786			C;Function:	
R;Mariyama, M.; Leinonen, A.; Mochizuki, T.; Tryggarson, K.; Reeder, S.T.			A;Description: minor structural component of extracellular basement membrane in kidney	
J. Biol. Chem., 269, 23013-23017, 1994			C;Superfamily: collagen alpha 1(IV) chain	
A;Title: Complete primary structure of the human alpha3(IV) collagen chain. Coexpression			C;Keywords: alternative splicing; basement membrane; cell binding; coiled coil; extra	
A;Reference number: A54763; MUID:94364994			F;1-28/Domain: signal sequence #status predicted <SIG>	
A;Accession: A54763			F;1-29-1670/Product: collagen alpha 3(IV) chain long splice form #status predicted <MA	
A;Molecule type: mRNA			F;1-29-428/Domain: amino-terminal nonhelical, NHI <NHI>	
A;Residues: 1-1670 <MARN>			F;7-9-1793/Region: cell attachment (R-G-D) motif	
A;Cross-references: GB:X80031; NID:9577563; PID:9577564			F;7-996-998/Region: cell attachment (R-G-D) motif	
A;Experimental source: kidney			F;1154-1156/Region: cell attachment (R-G-D) motif	
R;Turner, N.; Mason, P.J.; Brown, R.; Fox, M.; Povey, S.; Rees, A.; Pusey, C.D.			F;1306-1307/Region: cell attachment (R-G-D) motif	
J. Clin.-Invest., 89, 592-601, 1992			F;1345-1347/Region: cell attachment (R-G-D) motif	
A;Title: Molecular cloning of the human Goodpasture antigen demonstrates it to be the all			F;1432-1434/Region: cell attachment (R-G-D) motif	
A;Accession: A43928; MUID:92147878			F;1439-1670/Domain: carboxyl-terminal nonhelical, NCI <NCI>	
A;Reference number: A54763; MUID:92147878			F;1451-1551/Domain: collagen IV carboxyl-terminal repeat <CT>	
A;Accession: A54763			F;1561-1665/Domain: collagen IV carboxyl-terminal repeat <CT>	
A;Molecule type: mRNA			F;31-33, 39-41, 125-422, 476-479, 602, 722, 809-1387/Disulfide bonds: interchain #status predicted	
A;Residues: 1-1670 <MARN>			F;1553-1555/Region: carbohydrate (Asn) (covalent) #status predicted	
A;Cross-references: GB:X80031; NID:9577563; PID:9577564			F;1460-1548, 1493-1551/Disulfide bonds: (or 1460-1551, 1493-1548) #status predicted	
A;Experimental source: kidney			F;1505-1511, 1616-1622/Disulfide bonds: #status predicted	
R;Turner, N.; Mason, P.J.; Brown, R.; Fox, M.; Povey, S.; Rees, A.; Pusey, C.D.			F;1570-1662, 1604-1665/Disulfide bonds: (or 1570-1662, 1604-1665) #status predicted	
J. Clin.-Invest., 89, 592-601, 1992				
A;Title: Molecular cloning of the human Goodpasture antigen demonstrates it to be the all				
A;Accession: A43928; MUID:92147878				

Query Match 100.0%; Score 353; DB ID: Length 1670;
Best Local Similarity 100.0%; Pred. No. 5.5e-35;

RESULT 2

Qy 1 EFRASPFLCHGRCTCNYNSNSFWLASLNPERMFRKPKIPSTVKAGELEKISRCQVCM 60
 C;Species: *Mus musculus* (house mouse)
 C;Date: 15-Mar-1996 #sequence_revision 15-Mar-1996 #text_change 16-Feb-1997
 C;Accession: I48302; S4728

R;Miner, J.H.; Saines, J.R.
 J;Cell Biol. 127, 879-891, 1994
 A;Title: Collagen IV alpha₃, alpha₄, and alpha₅ chains in rodent basal laminae: sequence
 A;Accession: I48302
 A;Status: preliminary
 A;Molecule type: mRNA
 A;Residues: 1-246 <RES>
 A;Cross-references: EMBL:Z35166; NID:9535197; PID:9535198
 C;Superfamily: collagen alpha 1(IV) chain

Query Match 97.7%; Score 345; DB 2; Length 246;
 Best Local Similarity 95.3%; Pred. No. 7.8e-35; Indels 0; Gaps 0;
 Matches 61; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EFRASPFLCHGRCTCNYNSNSFWLASLNPERMFRKPKIPSTVKAGELEKISRCQVCM 60
 C;Accession: A39024; S20672; S17802; A15167; C39919; S13747; S20815
 C;Marrison, K.E.; Germino, G.G.; Reenders, S.T.
 J;Biol. Chem. 266, 34-39, 1991
 A;Title: Use of the polymerase chain reaction to clone and sequence a cDNA encoding the
 A;Reference number: A39024; MUID:91093146
 A;Accession: A39024
 A;Molecule type: mRNA
 A;Residues: 1-471 <MOR>
 A;Cross-references: EMBL:M63139; NID:9162886; PIDN:AA62708.1; PID:9162887
 R;Burkowski, R.J.; Langeveld, J.P.M.; Wieslander, J.; Hamilton, J.; Hudson, B.G.
 J;Biol. Chem. 262, 7874-7877, 1987
 A;Title: Localization of the Goodpasture epitope to a novel chain of basement membrane
 A;Reference number: S18412; MUID:8722219
 A;Accession: S20672
 A;Molecule type: protein
 A;Residues: 227-228, 'X', 230-244 <BUT>
 R;Saus, J.; Wieslander, J.; Langeveld, J.P.M.; Quinones, S.; Hudson, B.G.
 J;Biol. Chem. 263, 13374-13380, 1988
 A;Title: Identification of the Goodpasture antigen as the alpha₃(IV) chain of collagen
 A;Reference number: S17802; MUID:88330844
 A;Accession: S17802
 A;Molecule type: protein
 A;Residues: 227-228, 'X', 230-252, 'Y', 254 <SAU>
 R;Gunwar, S.; Saus, J.; Noelker, M.E.; Hudson, B.G.
 J;Biol. Chem. 265, 5466-5469, 1990
 A;Title: Glomerular basement membrane. Identification of a fourth chain, alpha₄, of type

RESULT 3

A39024
 collagen alpha 3(IV) chain - bovine (fragment)
 C;Species: *Bos primigenius taurus* (cattle)
 C;Date: 04-Dec-1992 #sequence_revision 04-Dec-1992 #text_change 13-Aug-1999
 C;Accession: A39024; S20672; S17802; A15167; C39919; S13747; S20815
 Qy 61 KKRH 64
 Db 243 KKRH 246

Query Match 91.5%; Score 323; DB 2; Length 471;
 Best Local Similarity 92.1%; Pred. No. 7.5e-32; Indels 0; Gaps 0;
 Matches 58; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 EFRASPFLCHGRCTCNYNSNSFWLASLNPERMFRKPKIPSTVKAGELEKISRCQVCM 60
 C;Accession: A39024; S20672; S17802; A15167; C39919; S13747; S20815
 C;Marrison, K.E.; Germino, G.G.; Reenders, S.T.
 J;Biol. Chem. 266, 34-39, 1991
 A;Title: Use of the polymerase chain reaction to clone and sequence a cDNA encoding the
 A;Reference number: A39024; MUID:91093146
 A;Accession: A39024
 A;Molecule type: mRNA
 A;Residues: 1-471 <MOR>
 A;Cross-references: EMBL:M63139; NID:9162886; PIDN:AA62708.1; PID:9162887
 R;Burkowski, R.J.; Langeveld, J.P.M.; Wieslander, J.; Hamilton, J.; Hudson, B.G.
 J;Biol. Chem. 262, 7874-7877, 1987
 A;Title: Localization of the Goodpasture epitope to a novel chain of basement membrane
 A;Reference number: S18412; MUID:8722219
 A;Accession: S20672
 A;Molecule type: protein
 A;Residues: 227-228, 'X', 230-244 <BUT>
 R;Saus, J.; Wieslander, J.; Langeveld, J.P.M.; Quinones, S.; Hudson, B.G.
 J;Biol. Chem. 263, 13374-13380, 1988
 A;Title: Identification of the Goodpasture antigen as the alpha₃(IV) chain of collagen
 A;Reference number: S17802; MUID:88330844
 A;Accession: S17802
 A;Molecule type: protein
 A;Residues: 227-228, 'X', 230-252, 'Y', 254 <SAU>
 R;Gunwar, S.; Saus, J.; Noelker, M.E.; Hudson, B.G.
 J;Biol. Chem. 265, 5466-5469, 1990
 A;Title: Glomerular basement membrane. Identification of a fourth chain, alpha₄, of type

Search completed: March 6, 2002, 06:54:49
 Job time: 40 sec

Copyright (c) 1993 - 2000 Compugen Ltd.

GenCore version 4.5

OM protein - protein search, using sw model

Run on: March 6, 2002, 07:05:09 ; Search time 18:75 Seconds

(without alignments)
125,149 Million cell updates/sec

Title: US-09-543-371-10_COPY_181-244
Perfect score: 353
Sequence: 1 EFRASPFLECHGRETQNYG.....KAGELEKLISRCQVCMKKRH 64

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 10059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 2

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 90%
Maximum Match 100%
Litting first 1000 summaries
Swissprot_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Database :

ALIGNMENTS

RESULT	Query No.	Score Match	Length	DB	ID	Description
1	353	100.0	1670	1	CA34_HUMAN	Q01955 homo sapien
2	323	91.5	471	1	CA34_BOVIN	Q28084 bos taurus

SEQUENCES

1
CA34_HUMAN STANDARD; PRT; 1670 AA.
AC 01955;
DT 01-OCT-1996 (Rel. 34, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE COLLAGEN ALPHA 3(IV) CHAIN PRECURSOR (GOODPASTURE ANTIGEN).
GN COL4A3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Etherilia; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TAXID:9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-KIDNEY.
RX MEDLINE=94368994; PubMed=8083201;
RA Miyama M., Leinonen A., Mochizuki T., Tryggvason K., Reenders S.T.;
RT "Complete primary structure of the human alpha 3(IV) collagen chain. Coexpression of the alpha 3(IV) and alpha 4(IV) collagen chains in human tissues.";
RL J. Biol. Chem. 269:23013-23017(1994).
RN [2]
RP REVISIONS.
RA Leinonen A.;
RA Submitted (Oct-1998) to the EMBL/GenBank/DDBJ databases.

2
RA SEQUENCE OF 1453-1670 FROM N.A.
RX MEDLINE=91353570; PubMed=1882840;
RA Morrison K.E., Marinova M., Yang-Feng T.L., Reenders S.T.;
RT "Sequence and localization of a partial cDNA encoding the human alpha 3 chain of type IV collagen";
RL Am. J. Hum. Genet. 49:545-554 (1991).
RN [5]
RP SEQUENCE OF 1331-1670 FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=92147878; PubMed=173749;
RA Turner N., Mason P.J., Brown R., Fox M., Povey S., Rees A.,
RA Pusey C.D.;
RT "Molecular cloning of the human Goodpasture antigen demonstrates it to be the alpha 3 chain of type IV collagen.";
RL J. Clin. Invest. 89:592-601 (1992).
RN [6]
RP SEQUENCE OF 1644-1670 FROM N.A.
RC TISSUE=Kidney;
RA Ding J.;
RL Submitted (Jan-1993) to the EMBL/GenBank/DDBJ databases.
RP SEQUENCE OF 1439-1670, AND ALTERNATIVE SPlicing.
RX MEDLINE=94124597; PubMed=8294492;
RA Feng L., Xia Y., Wilson C.B.;
RT "Alternative splicing of the NCI domain of the human alpha 3(IV) collagen gene. Differential expression of mRNA transcripts that predict three protein variants with distinct carboxyl regions.";
RT J. Biol. Chem. 269:2342-2348 (1994).
RN [8]
RP SEQUENCE OF 1-29 FROM N.A.
RX MEDLINE=98196854; PubMed=953706;
RA Momota R., Sugimoto M., Oohashi T., Kigasawa K., Yoshioka H.,
RA Ninomaya Y.;
RT "Two genes, COL4A3 and COL4A4 coding for the human alpha3(IV) and alpha4(IV) collagen chains are arranged head-to-head on chromosome 2q36.;"
RL FEBS Lett. 424:11-16 (1998).
RN [9]
RP ALTERNATIVE SPlicing.
RX MEDLINE=93280184; PubMed=8505332;
RA Bernal D., Quinones S., Saus J.;
RT "The human mRNA encoding the Goodpasture antigen is alternatively spliced.";
RL J. Biol. Chem. 268:12090-12094 (1993).
RN [10]
RP VARIANT PRO-1474.
RX MEDLINE=95078827; PubMed=7987301;
RA Lemmink H.H., Mochizuki T., van den Heuvel L.P.W.J., Schroeder C.H.,
RA Barrientos A., Monnens L.A.H., van Oost B.A., Brunner H.G.,
RA Reenders S.T., Smeets H.J.M.;
RT "Mutations in the type IV collagen alpha 3 (COL4A3) gene in autosomal recessive Alport syndrome";
RL Hum. Mol. Genet. 3:1269-1273 (1994).
RN [11]
RP VARIANTS AS AND VARIANTS.
RX MEDLINE=21064696; PubMed=11134255;
RA Heidet L., Arrendel C., Forestier L., Cohen-Solal L., Mollet G.,
RA Gutierrez B., Stavrou C., Gubler M.C., Antignac C.;
RT "Structure of the human type IV collagen gene COL4A3 and mutations in autosomal Alport syndrome";
RL J. Am. Soc. Nephrol. 12:97-106 (2001).
CC -!- FUNCTION: TYPE IV COLLAGEN IS THE MAJOR STRUCTURAL COMPONENT OF GLOMERULAR BASEMENT MEMBRANES (GBM), FORMING A 'CHICKEN WIRE' SEQUENCE OF 1386-1670 FROM N.A., AND PARTIAL SEQUENCE. RX MEDLINE=93015826; PubMed=1400291;
RA Quinones S., Bernal D., Garcia-Sogo M., Elena S.F., Saus J.;
RT "Exon/inton structure of the human alpha 3(IV) gene encompassing the Goodpasture antigen (alpha 3(IV)NC1). Identification of a potentially antigenic region at the triple helix/NC1 domain junction.";
RT J. Biol. Chem. 267:19780-19784 (1992).
RN [4]

CC MESHWORK TOGETHER WITH LAMININS, PROTEOGLYCANS AND ENACTIN/ NIDROGEN.

-!- SUBUNIT: THERE ARE SIX TYPE IV COLLAGEN ISOFORMS; ALPHA 1(IV)- ALPHAS 6(IV), EACH OF WHICH CAN FORM A TRIPLE HELIX STRUCTURE WITH 2 OTHER CHAINS TO GENERATE TYPE IV COLLAGEN NETWORK.

-!- SUBCELLULAR LOCATION: CELL SURFACE (POTENTIAL).

-!- ALTERNATIVE PRODUCTS: AT LEAST 3 ISOFORMS; 1 (SHOWN HERE), 2/V AND 3/L5; ARE PRODUCED BY ALTERNATIVE SPlicing. THEY DIFFER IN THEIR C-TERMINAL NC1 DOMAINS.

-!- TISSUE SPECIFICITY: ALPHA 3 AND ALPHA 4 TYPE IV COLLAGENS ARE COLOCALIZED AND PRESENT ONLY IN BASEMENT MEMBRANES OF KIDNEY, EYE, COCHLEA, LUNG AND BRAIN.

-!- DOMAIN: ALPHA CHAINS OF TYPE IV COLLAGEN HAVE A NONCOLLAGENOUS DOMAIN (NC1) AT THEIR C TERMINUS, FREQUENT INTERRUPTIONS OF THE G-X-Y REPEATS IN THE LONG CENTRAL TRIPLE-HELICAL DOMAIN (WHICH MAY CAUSE FLEXIBILITY IN THE TRIPLE HELIX), AND A SHORT N-TERMINAL TRIPLE-HELICAL 7S DOMAIN.

-!- PTM: PROLINES AT THE THIRD POSITION OF THE TRIPETIDE REPEATING UNIT (G-X-Y) ARE HYDROXYLATED IN SOME OR ALL OF THE CHAINS.

-!- PTM: THE ALTERNATIVE SPliced FORM V CONTAINS AN ADDITIONAL N-LINKED GLYCOSYLATION SITE.

-!- PTM: TYPE IV COLLAGENS CONTAIN NUMEROUS CYSTEINE RESIDUES WHICH ARE INVOLVED IN INTRAMOLECULAR DISULFIDE BONDING. 12 OF THESE, LOCATED IN THE NC1 DOMAIN, ARE CONSERVED IN ALL KNOWN TYPE IV COLLAGENS.

-!- DISPERSE: ANTIBODIES AGAINST THE NC1 DOMAIN OF ALPHA3(IV) MEDIATE THE HUMAN AUTOIMMUNE DISEASE, GOODPASTURE SYNDROME, WHICH IS CHARACTERIZED BY HEMATURIA AND PULMONARY HEMORRHAGE.

-!- DISPERSE: DEFECTS IN COLA3 ARE ASSOCIATED WITH THE TYPE I AUTOSOMAL RECESSIVE FORM OF ALPORT SYNDROME, AN HEREDITARY GLOMERULONEPHROPATHY CHARACTERIZED BY PROGRESSIVE RENAL FAILURE, HEMATURIA AND DEAFNESS. THE RECESSIVE FORM OCCURS EQUALLY BETWEEN MALES AND FEMALES.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

FT	SITE	791	793	CELL ATTACHMENT SITE (POTENTIAL).
FT	SITE	996	998	CELL ATTACHMENT SITE (POTENTIAL).
FT	SITE	1154	1156	CELL ATTACHMENT SITE (POTENTIAL).
FT	SITE	1306	1308	CELL ATTACHMENT SITE (POTENTIAL).
FT	SITE	1345	1347	CELL ATTACHMENT SITE (POTENTIAL).
FT	SITE	1432	1434	CELL ATTACHMENT SITE (POTENTIAL).
FT	CARBOHYD	253	253	N-LINKED (GLCNAC, . .) (POTENTIAL).
FT	MOD_RES	1435	1435	PHOSPHORYLATION (BY SIMILARITY).
FT	MOD_RES	1437	1437	PHOSPHORYLATION (BY SIMILARITY).
FT	DISLFID	1460	1551	OR 1548 (BY SIMILARITY).
FT	DISLFID	1493	1548	OR 1551 (BY SIMILARITY).
FT	DISLFID	1505	1511	BY SIMILARITY.
FT	DISLFID	1570	1665	OR 1662 (BY SIMILARITY).
FT	DISLFID	1604	1662	OR 1665 (BY SIMILARITY).
FT	VARSPLIC	1586	1670	FTSAGSEGTCQALASPGLCLEEFRASPFLCHGRGTCTNYSNSFWLASINPERMFRRKIPPSIVKAGELEKTIISRCVCMK NSYSNLASINPERMFRRKIPPSIVKAGELEKTIISRCVCMK KRH -> KAVSINCEWSGKRKNSLSGVWEEKRKKTA ELVFFLNKVKMTHAVI (IN ISOFORM 2).
FT	VARSPLIC	1488	1670	GTLSSCLQRPTTMPLFCNNDNVNPASNDRDSYWLSPAL GPMNNATGRLAPYLSRCTVCGGPATLAVISOTDPP CPHGHTSLWKGFSITMTSAGSEGICQALASPGLCLEEFA SPFLCHGRGTCTNYSNSFWLASINPERMFRRKIPPSIVK AGELEKTIISRCVCMK KRH -> DALFWVYLRS (IN ISOFORM 3).
FT	VARIANT	43	43	G -> R.
FT	VARIANT	162	162	/FTId=VAR_011202.
FT	VARIANT	297	297	G -> E (IN AST).
FT	VARIANT	61	KKRH	G -> E (IN AST).
Db	1667	KKRH	1670	
Query	Match	100.0%	Score 353; DB 1; Length 1670;	
Query	Best Local Similarity	100.0%	Pred. No. 2.7e-36;	
Query	Matches	64;	Conservative 0;	
Query	Mismatches	0;	Mismatches 0;	
Query	Indels	0;	Indels 0;	
Query	Gaps	0;	Gaps 0;	
Db	1607	EFRASPFLECHGRGTCTNYSNSFWLASINPERMFRRKIPPSIVKAGELEKTIISRCVCM	60	
Db	1666	EFRAASPFLECHGRGTCTNYSNSFWLASINPERMFRRKIPPSIVKAGELEKTIISRCVCM	1666	
RESULT	2			
CA34_BOVIN				
AC	CA34_BOVIN		STANDARD;	PRT;
DT	028084;			471 AA.
DT	01-Nov-1997	(Rel. 35, Created)		
DT	01-Nov-1997	(Rel. 35, last sequence update)		
DT	01-Nov-1997	(Rel. 35, Last annotation update)		
DE	COLLAGEN ALPHA 3(IV) CHAIN (FRAGMENT).			
GN	COL4A3.			
OS	Bos taurus (Bovine).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;			
OC	Bovidae; Bovinae; Bos.			
RN	[1].			
RP	SEQUENCE FROM N.A.			
RC	TISSUE= lens;			
RX	Medline=91033146; PubMed=1985905;			
RA	Morrison R.E., Gemino G.G., Readers S.T.,			
RT	Use of the polymerase chain reaction to clone and sequence a cDNA encoding the bovine alpha 3 chain of type IV collagen.;			
RL	J. Biol. Chem. 266:34 39 (1991).			
RN	[2].			
RP	SEQUENCE OF 227-258.			
RP	TISSUE=Kidney;			
RX	Medline=90202779; PubMed=2318822;			
RA	Gunwar S., Saus J., Moelker M.E., Hudson B.G.;			
RT	"Glomerular basement membrane. Identification of a fourth chain, alpha 4, of type IV collagen.";			
RT	(By SIMILARITY).			

RL J. Biol. Chem. 265:5466-5469(1990).
 RN [3]
 RP SEQENCE OF 227-254.
 RX MEDLINE=88310844; PubMed=3417661;
 RA Saus J., Wieslander J., Langeveld J.P.M., Quinones S., Hudson B.G.;
 RT "Identification of the Goodpasture antigen as the alpha 3(IV) chain
 of collagen IV.";
 J. Biol. Chem. 263:13374-13380(1988).
 [4]
 RT SEQUENCE OF 227-244.
 RX MEDLINE=87222419; PubMed=2438283;
 RA Butkowksi R.J., Langeveld J.P.M., Wieslander J., Hamilton J.,
 RA Hudson B.G.;
 RT Localization of the Goodpasture epitope to a novel chain of basement
 membrane collagen.;"
 RL J. Biol. Chem. 262:7874-7877(1987).
 CC FUNCTION: TYPE IV COLLAGEN IS THE MAJOR STRUCTURAL COMPONENT OF
 GLOMERULAR BASEMENT MEMBRANES (GBM), FORMING A 'CHICKEN-WIRE'
 MESHWORK TOGETHER WITH LAMININS, PROFOGlyCANS AND ENTACTIN/
 CC NIDGEN.
 CC -!- SUBUNIT: THERE ARE SIX TYPE IV COLLAGEN ISOFORMS, ALPHA 1(IV)-
 CC ALPHA 6(IV), EACH OF WHICH CAN FORM A TRIPLE HELIX STRUCTURE
 CC WITH 2 OTHER CHAINS TO GENERATE TYPE IV COLLAGEN NETWORK.
 CC -!- SUBCELLULAR LOCATION: CELL SURFACE (POTENTIAL).
 CC -!- DOMAIN: ALPHA CHAINS OF TYPE IV COLLAGEN HAVE A NONCOLLAGENOUS
 CC DOMAIN (NC1) AT THEIR C-TERMINUS, FREQUENT INTERRUPTIONS OF THE
 CC G-X-Y REPEATS IN THE LONG CENTRAL TRIPLE-HELICAL DOMAIN (WHICH MAY
 CC CAUSE FLEXIBILITY IN THE TRIPLE HELIX), AND A SHORT N-TERMINAL
 CC TRIPLE-HELICAL 7S DOMAIN.
 CC -!- PTM: PROLINES AT THE THIRD POSITION OF THE TRIPETIDE REPEATING
 CC UNIT (G-X-Y) ARE HYDROXYLATED IN SOME OR ALL OF THE CHAINS.
 CC -!- PTM: TYPE IV COLLAGENS CONTAIN NUMEROUS CYSTEINE RESIDUES WHICH
 CC ARE INVOLVED IN INTER- AND INTRAMOLECULAR DISULFIDE BONDING. 12 OF
 CC THESE, LOCATED IN THE NC1 DOMAIN, ARE CONSERVED IN ALL KNOWN TYPE
 CC IV COLLAGENS.
 CC -!- SIMILARITY: TO OTHER TYPE IV COLLAGENS.
 CC -----
 CC THIS SWISS-PROT ENTRY IS COPYRIGHT. IT IS PRODUCED THROUGH A COLLABORATION
 CC BETWEEN THE SWISS INSTITUTE OF BIOINFORMATICS AND THE EMBL OUTSTATION -
 CC THE EUROPEAN BIOINFORMATICS INSTITUTE. THERE ARE NO RESTRICTIONS ON ITS
 CC USE BY NON-PROFIT INSTITUTIONS AS LONG AS ITS CONTENT IS IN NO WAY
 CC MODIFIED AND THIS STATEMENT IS NOT REMOVED. USAGE BY AND FOR COMMERCIAL
 CC ENTITIES REQUIRES A LICENSE AGREEMENT (SEE [HTTP://WWW.IB-SIB.CH/ANNOUNCE/](http://www.ib-sib.ch/announce/)
 CC OR SEND AN EMAIL TO license@sib.ch).
 CC -----
 CC EMBL; M63139; ARX62708.1; .
 CC DR InterPro; IPR001442; C4;
 CC DR InterPro; IPR00087; Collagen.
 CC DR Pfam; PF01413; C4; 2.
 CC DR ProDom; PD003923; C4; 2.
 CC SMART; SM00111; C4; 2.
 KW Extracellular matrix; Connective tissue; Repeat; Hydroxylation;
 KW Glycoprotein; Basement membrane; Collagen; Cell adhesion.
 FT NON_TER 1 1
 FT DOMAIN <1 238 TRIPLE-HELICAL REGION.
 FT DOMAIN 239 471 NONHELICAL REGION (NC1).
 FT SITE 106 108 CELL ATTACHMENT SITE (POTENTIAL).
 FT MOD_RES 232 232 HYDROXYLATION.
 FT MOD_RES 238 238 HYDROXYLATION.
 FT DISULFD 261 352 OR 349 (BY SIMILARITY).
 FT DISULFD 294 349 OR 352 (BY SIMILARITY).
 FT DISULFD 306 312 BY SIMILARITY.
 FT DISULFD 371 466 OR 463 (BY SIMILARITY).
 FT DISULFD 405 463 OR 466 (BY SIMILARITY).
 FT DISULFD 417 423 BY SIMILARITY.
 FT DISULFD 253 S -> Y (IN REF. 3).
 SQ SEQUENCE 471 AA; 47585 MW; C03B66F14E7008DE CRC64;

Query Match 91.5%; Score 323; DB 1; Length 471;
 Best Local Similarity 92.1%; Pred. No. 4e-33; 3; Mismatches 58;
 Matches 58; Conservative 2; Indels 0; Gaps 0;

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on:

March 6, 2002, 07:04:44 ; Search time 50.56 Seconds

(without alignments)

185.155 Million cell updates/sec

Title: US-09-543-371-10_COPY_181_244
Perfect score: 353
Sequence: 1 EFRASPFLECHGRGTCNYYS..... KAGELEKILISRCQVCMKKRH 64

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

T number of hits satisfying chosen parameters: 7
Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 90%
Maximum Match 100%
Listing first 1000 summaries

Database : SPTREMBL17:
1: sp_archaea:
2: sp_bacteria:
3: sp_fungi:
4: sp_human:
5: sp_invertebrate:
6: sp_mammal:
7: sp_mhc:
8: sp_organelle:
9: sp_phage:
10: sp_plant:
11: sp_rodent:
12: sp_virus:
13: sp_vertebrate:
14: sp_unclassified:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	353	100.0	212	6 Q28512	Q28512 macaca mula
2	353	100.0	245	4 Q9NYC4	Q9NYC4 homo sapien
3	353	100.0	1670	4 Q9BQ72	Q9BQ72 homo sapien
4	348	98.6	230	11 Q63122	Q63122 rattus norv
5	348	98.6	246	11 Q61435	Q61435 mus musculus
6	348	98.6	1659	11 Q9qzs0	Q9qzs0 mus musculus
7	331	93.8	212	6 Q28567	Q28567 ovis aries

ALIGNMENTS

RESULT	1
ID	Q28512
AC	Q28512;
DT	01-NOV-1996 (TREMBREL_01, Created)
DT	01-NOV-1996 (TREMBREL_01, Last sequence update)
DT	01-JUN-2001 (TREMBREL_17, Last annotation update)

PRELIMINARY; PRT; 212 AA.

DE	ALPHA-3 TYPE IV COLLAGEN (FRAGMENT).
GN	COL4A3.
OS	Macaca mulatta (Rhesus macaque).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae; Cercopithecinae; Macaca.
OC	NBII_TaxID=9544;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	TISSUE-RENAL CORTEX;
RA	Turner A.N., Ryan J.J., Derry C.J., Cashman S.J., Katbamna I., Mason P.J., Pusey C.D.; Submitted (MAR 1996) to the EMBL/GenBank/DBJ databases.
RL	EMBL; 147280; AAA1861.1; -.
DR	INTERPRO; IPR00504; RRM.
DR	INTERPRO; IPR00142; C4.
DR	Pfam; PF01413; C4; 2.
DR	PRODOM; PD003923; C4; 2.
DR	SMART; SM00111; C4; 2.
FT	NON_TER 212 212 212 AA; 23469 MW; 4BC574A64E357E64 CRC64;
FT	SEQUENCE 212 AA; 23469 MW; 4BC574A64E357E64 CRC64;
Query	Match 100.0%; Score 353; DB 6; Length 212;
Best Local Similarity	100.0%; Pred. No. 1.8e-36; 0; Mismatches 0; Indels 0; Gaps 0;
Matches	64; Conservative 0; MisMatches 0; Indels 0; Gaps 0;
QY	1 EFRASPFLECHGRGTCNYYSNSYFWLASNPERMFRKPIPSVKAGELEKILISRCQVCM 60
Db	149 EFRASPFLECHGRGTCNYYSNSYFWLASNPERMFRKPIPSVKAGELEKILISRCQVCM 208
QY	61 KKRH 64
Db	209 KKRH 212
RESULT	2
ID	Q9NYC4
AC	Q9NYC4
DT	01-OCT-2000 (TREMBREL_15, Created)
DT	01-OCT-2000 (TREMBREL_15, Last sequence update)
DE	TUMSTATIN (FRAGMENT).
OS	Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID	=9606;
RN	[1]
RP	SEQUENCE FROM N.A.
RA	Mashima Y., Colorado P.C., Torre A., Holthaus K.A., Grunkemeyer J.A., Erickson M.D., Hooper H., Xiao Y., Stillman T.E., Kalluri R.; RT "Distinct anti-tumor properties of a type IV collagen domain derived from basement membrane." J. Biol. Chem. 2000; 275(20):14000-14005.
RL	EMBL; AR25851; AAF72632; 1; -.
DR	INTERPRO; IPR00142; C4.
DR	INTERPRO; IPR00504; RRM.
DR	Pfam; PF01413; C4; 2.
DR	PRODOM; PD003923; C4; 2.
DR	SMART; SM00111; C4; 2.
DR	PROSITE; PS00030; RRM_RNP_1; UNKNOWN_1.
FT	NON_TER 1 245 AA; 26952 MW; 1EE5028354D9A57D CRC64;
FT	SEQUENCE 1 245 AA; 26952 MW; 1EE5028354D9A57D CRC64;
Query	Match 100.0%; Score 353; DB 4; Length 245;
Best Local Similarity	100.0%; Pred. No. 2.1e-36; 0; Mismatches 0; Indels 0; Gaps 0;
Matches	64; Conservative 0; MisMatches 0; Indels 0; Gaps 0;
QY	1 EFRASPFLECHGRGTCNYYSNSYFWLASNPERMFRKPIPSVKAGELEKILISRCQVCM 60

Db	182	EFRASPFLECHGRGTCNYYSNSYFWLASLNPERMFRKIPSTVKAGELEKIIISRCQVCM	241	DR EMBL; AJ288531; CAC36101..1; JOINED.
Qy	61	KKRH 64		DR EMBL; AJ288333; CAC36101..1; JOINED.
Db	242	KKRH 245		DR EMBL; AJ288534; CAC36101..1; JOINED.
RESULT	3			DR EMBL; AJ288535; CAC36101..1; JOINED.
ID	Q9BQT2	PRELIMINARY;	PRT;	DR EMBL; AJ288536; CAC36101..1; JOINED.
AC	Q9BQT2;			DR EMBL; AJ288537; CAC36101..1; JOINED.
DT	01-JUN-2001	(TRIMBLEL. 17, Created)		
DT	01-JUN-2001	(TRIMBLEL. 17, Last sequence update)		
DE	ALPHA3 TYPE IV COLLAGEN.			
GN	COLA3?			
OS	Homo sapiens (Human)			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.			
NN	[1]			
RE	SEQUENCE FROM N_A.			
RX	MEDLINE=21064690; PubMed=11134255;			
RA	Heidet L., Arrendell C., Forrestier L., Cohen-Solal L., Mollet G., Gutierrez B., Stavrou C., Gubler M.C., Antignac C.,			
RT	"Structure of the human type IV collagen gene COLA3 and mutations in autosomal Alport syndrome."			
RL	J. Am. Soc. Nephrol. 12:97-106(2001).			
DR	EMBL; AJ288487; CAC36101..1; -			
DR	EMBL; AJ288488; CAC36101..1; JOINED.			
DR	EMBL; AJ288489; CAC36101..1; JOINED.			
DR	EMBL; AJ288490; CAC36101..1; JOINED.			
DR	EMBL; AJ288491; CAC36101..1; JOINED.			
DR	EMBL; AJ288492; CAC36101..1; JOINED.			
DR	EMBL; AJ288493; CAC36101..1; JOINED.			
DR	EMBL; AJ288494; CAC36101..1; JOINED.			
DR	EMBL; AJ288495; CAC36101..1; JOINED.			
DR	EMBL; AJ288496; CAC36101..1; JOINED.			
DR	EMBL; AJ288497; CAC36101..1; JOINED.			
DR	EMBL; AJ288498; CAC36101..1; JOINED.			
DR	EMBL; AJ288499; CAC36101..1; JOINED.			
DR	EMBL; AJ288500; CAC36101..1; JOINED.			
DR	EMBL; AJ288501; CAC36101..1; JOINED.			
DR	EMBL; AJ288502; CAC36101..1; JOINED.			
DR	EMBL; AJ288503; CAC36101..1; JOINED.			
DR	EMBL; AJ288504; CAC36101..1; JOINED.			
EMBL	AJ288505; CAC36101..1; JOINED.			
EMBL	AJ288506; CAC36101..1; JOINED.			
EMBL	AJ288507; CAC36101..1; JOINED.			
DR	EMBL; AJ288508; CAC36101..1; JOINED.			
DR	EMBL; AJ288509; CAC36101..1; JOINED.			
DR	EMBL; AJ288510; CAC36101..1; JOINED.			
DR	EMBL; AJ288511; CAC36101..1; JOINED.			
DR	EMBL; AJ288512; CAC36101..1; JOINED.			
DR	EMBL; AJ288513; CAC36101..1; JOINED.			
DR	EMBL; AJ288514; CAC36101..1; JOINED.			
DR	EMBL; AJ288515; CAC36101..1; JOINED.			
DR	EMBL; AJ288516; CAC36101..1; JOINED.			
DR	EMBL; AJ288517; CAC36101..1; JOINED.			
DR	EMBL; AJ288518; CAC36101..1; JOINED.			
DR	EMBL; AJ288519; CAC36101..1; JOINED.			
DR	EMBL; AJ288520; CAC36101..1; JOINED.			
DR	EMBL; AJ288521; CAC36101..1; JOINED.			
DR	EMBL; AJ288522; CAC36101..1; JOINED.			
DR	EMBL; AJ288523; CAC36101..1; JOINED.			
DR	EMBL; AJ288524; CAC36101..1; JOINED.			
DR	EMBL; AJ288525; CAC36101..1; JOINED.			
DR	EMBL; AJ288526; CAC36101..1; JOINED.			
DR	EMBL; AJ288527; CAC36101..1; JOINED.			
DR	EMBL; AJ288528; CAC36101..1; JOINED.			
DR	EMBL; AJ288529; CAC36101..1; JOINED.			
DR	EMBL; AJ288530; CAC36101..1; JOINED.			
RESULT	5			
ID	Q61435	PRELIMINARY;	PRT;	DR EMBL; AJ288535; CAC36101..1; JOINED.
AC	Q61435;			DR EMBL; AJ288536; CAC36101..1; JOINED.
SQ	SEQUENCE 1670 AA; 161899 MW; FA7BE4914CA0A6F6 CRC64;			DR EMBL; AJ288537; CAC36101..1; JOINED.
RESULT	4			DR EMBL; AJ288538; CAC36101..1; JOINED.
ID	Q63122	PRELIMINARY;	PRT;	DR EMBL; AJ288539; CAC36101..1; JOINED.
AC	Q63122			DR EMBL; AJ288539; CAC36101..1; JOINED.
DT	01-MAY-1999 (TRIMBLEL. 01, Created)			
DT	01-JUN-2001 (TRIMBLEL. 10, Last sequence update)			
DE	ALPHA-3 TYPE IV COLLAGEN (FRAGMENT).			
GN	COLA3?			
OS	Rattus norvegicus (Rat)			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			
OX	[1]			
RN	NCBI-TAXID=10116;			
RP	SEQUENCE FROM N_A.			
RC	STRAIN=SPRAGUE-DAWLEY; TISSUE=KIDNEY CORTEX;			
RA	Turner N., Ryan J.J., Deery C.J., Cashman S.J., Katbana I., Mason P.J., Pusey C.D.; Submitted (MAR-1999) to the EMBL/GenBank/DDBJ databases.			
RL	R. L. Turner N., Ryan J.J., Deery C.J., Cashman S.J., Katbana I., Mason P.J., Pusey C.D.; Submitted (MAR-1999) to the EMBL/GenBank/DDBJ databases.			
DR	EMBL; L47781; AAB72338; -			
DR	InterPro; IPR000504; RRM.			
DR	InterPro; IPR001442; C4.			
DR	PFam; PF04143; C4; 2.			
DR	ProdDom; PD00323; C4; 2.			
DR	PROSITE; PS0030; RRM_RNP_1; UNKNOWN_1.			
DR	SMART; SM0011; C4; 2.			
FT	NON_TER 1			
FT	NON_TER 230			
FT	SEQUENCE 230 AA; 25398 MW; 29549E25314CC056 CRC64;			
Query	Match 98 %	Score 348;	DB 11;	Length 230;
Best Local Similarity 96.9%	Pred. No. 8.5e-36;	0;	Indels 0;	Gaps 0;
Matches 62;	Conservative 2;	Mismatches 0;		
Qy	1 EFRASPFLECHGRGTCNYYSNSYFWLASLNPERMFRKIPSTVKAGELEKIIISRCQVCM 60			
Db	167 EFRASPFLECHGRGTCNYYSNSYFWLASLNPERMFRKIPSTVKAGELEKIIISRCQVCM 226			
Qy	61 KKRH 64			
Db	227 KKRH 230			
Query Match 98 %	Score 348;	DB 11;	Length 230;	
Best Local Similarity 96.9%	Pred. No. 8.5e-36;	0;	Indels 0;	Gaps 0;
Matches 62;	Conservative 2;	Mismatches 0;		
Qy	1 EFRASPFLECHGRGTCNYYSNSYFWLASLNPERMFRKIPSTVKAGELEKIIISRCQVCM 60			
Db	167 EFRASPFLECHGRGTCNYYSNSYFWLASLNPERMFRKIPSTVKAGELEKIIISRCQVCM 226			
Qy	61 KKRH 64			
AC	Q61435	PRELIMINARY;	PRT;	DR EMBL; AJ288535; CAC36101..1; JOINED.
ID	Q61435;			DR EMBL; AJ288536; CAC36101..1; JOINED.
AC	Q61435;			DR EMBL; AJ288537; CAC36101..1; JOINED.

DT DT 01-NOV-1996 (TREMBlrel. 01, Created)
 01-JUN-2001 (TREMBlrel. 01, Last annotation update)
 PROCOLLAGEN TYPE IV, ALPHA 3 (COLLAGEN IV ALPHA 3 CHAIN) (FRAGMENT).
 COL4A3.

GN OS
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathii; Muridae; Murinae; Mus.
 Mus musculus (Mouse).

OC OC
 Mammalia; Eutheria; Rodentia; Sciurognathii; Muridae; Murinae; Mus.
 Mus musculus (Mouse);
 Mammalia; Eutheria; Rodentia; Sciurognathii; Muridae; Murinae; Mus.

NCBI_TAXID=10090;
 [1] DR
 SEQUENCE FROM N.A.
 STRAIN=BALB/C;
 MEDLINE=95050957; PubMed=7962065;
 Miner J.H.; Sanes J.R.;
 "Collagen IV alpha 3, alpha 4, and alpha 5 chains in rodent basal
 laminae: sequence, distribution, association with laminins, and
 developmental switches."; J. Cell Biol. 127:879-891 (1994).

RP RP
 SEQUENCE FROM N.A.
 STRAIN=BALB/C;
 Miner J.H.;
 RL DR
 Submitted (JUL-1994) to the EMBL/GenBank/DDJB databases.
 [3] RN
 SEQUENCE FROM N.A.
 STRAIN=BALB/C;
 Miner J.H.;
 RL DR
 Submitted (MAY-1996) to the EMBL/GenBank/DDJB databases.
 EMBL; Z35166; CAR84529.1; -.
 MGD; MGI:10468; Col4a3.
 DR DR
 InterPro; IPR01442; C4.
 InterPro; IPR0143; C4.
 InterPro; IPR00504; RRM.
 InterPro; IPR01413; C4; 2.
 Pfam; PF0143; C4; 2.
 Prodom; PDO03923; C4; 2.
 SMART; SM00111; C4; 2.
 PROSITE; PS00030; RRM_RNP_1; UNKNOWN_1.
 FT FT
 NON_TER 1 I
 SQ SEQUENCE 246 AA; 26993 MW; A9B5434F5B36F324 CRC64;

Query Match 98.6%; Score 348; DB 11; Length 246;
 Best Local Similarity 96.9%; Pred. No. 9; 1e-36; 0; Gaps 0;
 Matches 62; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy Qy 1 EFRASPELECHGRGTTCNVYNSNSFWLASLNPERMFRKPIPSVKAGLEKITSRCQVCM 60
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
 Db 1606 EFRASPELECHGRGTTCNVYNSNSFWLASLNPERMFRKPIPSVKAGLEKITSRCQVCM 1665

Qy Qy 61 KRRH 64
 |||||:
 Db 1666 KRRH 1669

RESULT 7
 ID Q88567 PRELIMINARY; PRT; 212 AA.
 AC Q88567;
 DT 01-NOV-1996 (TREMBlrel. 01,保守性)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE ALPHA-3 TYPE IV COLLAGEN (FRAGMENT).
 GN OVIS aries (Sheep).
 OS Ovis aries (Sheep).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 Bovidae; Caprinae; Ovis.
 OC NCBI_TAXID=9940;
 RN {1}
 RP SEQUENCE FROM N.A.
 RC TISSUE-RENAL CORTEX;
 RA Turner A.N.; Ryan J.J.; Derry C.J.; Cashman S.J.; Katbamna L.;
 RA Mason P.J.; Pusey C.D.;
 RL DR
 Submitted (MAR-1996) to the EMBL/GenBank/DDJB databases.
 EMBL; I47282; AA91904.1; -.
 DR InterPro; IPR00504; RRM.
 DR InterPro; IPR01442; C4.
 DR Pfam; PF0143; C4; 2.
 DR Prodom; PDO03923; C4; 2.
 DR PROSITE; PS00030; RRM_RNP_1; UNKNOWN_1.
 DR SMART; SM00111; C4; 2.
 FT FT
 NON_TER 1 I
 SQ SEQUENCE 212 AA; 23417 MW; 0F5839FCB81BDD8C CRC64;

Query Match 93.8%; Score 331; DB 6; Length 212;
 Best Local Similarity 92.1%; Pred. No. 1.1e-33; 0; Gaps 0;
 Matches 58; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Qy Qy 1 EFRASPELECHGRGTTCNVYNSNSFWLASLNPERMFRKPIPSVKAGLEKITSRCQVCM 60
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:
 Db 149 EFRASPELECHGRGTTCNVYNSNSFWLASLNPERMFRKPIPSVKAGLEKITSRCQVCM 208

Qy Qy 61 KRR 63
 |||||:
 Db 209 KRR 211

Wed Mar 6 07:07:42 2002

us-09-543-371-10_copy_181_244.rspt

Page 4

Search completed: March 6, 2002, 07:04:44
Job time: 590 sec

KW Goodpasture syndrome; type IV collagen; alpha3 chain; human.
 XX
 OS Homo sapiens.
 XX
 PN US6007980-A.
 XX
 PD 28-DEC-1999.
 XX
 PR 07-OCT-1998; 98US-0167364.
 XX
 PR 30-NOV-1990; 90US-0621091.
 XX
 PR 07-MAR-1995; 95US-039889.
 XX
 (UNIV) UNIV KANSAS MEDICAL CENT.
 PA (UYYA) UNIV YALE.
 XX
 PI Hudson BG, Reeder ST, Morrison KE;
 XX
 DR WPI; 2000-096371/08.
 N-PSDB; AAZ46729.
 PT Diagnosing and treating Goodpasture syndrome using a peptide derived
 from type IV collagen -
 XX
 PS Disclosure; Columns 23-26; 26pp; English.
 XX
 CC The invention provides a method of detecting Goodpasture antibodies
 in the fluid of a patient by contacting it with a peptide comprising at
 most 218 amino acids of the human alpha3 chain type IV collagen that
 contains the fragment shown in AAV56785. The methods are useful for the
 diagnosis and treatment of Goodpasture syndrome. The present sequence
 represents the carboxy terminal noncollagenous domain of the human
 alpha3 chain of type IV collagen.
 XX
 Sequence 218 AA;
 SQ
 Query Match 100.0%; Score 353; DB 21; Length 218;
 Best Local Similarity 100.0%; Pred. No. 9e-39;
 Matches 64; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 EFRASPFLRCHGRGTCNYYSNSYNSFWLAINPERMFRRKPIPSVKA
 Db 155 efraspflechgrgtnyyssnsysfwlasinpermfrkpiptvka
 QY 61 KKRH 64
 YY 215 kkrh 218
 DE
 RESULT 3
 AAY67942
 ID AAY67942 standard; Protein; 245 AA.
 AC
 XX
 AC AAY67942;
 DT 03-APR-2000 (first entry)
 XX
 DE Human type IV collagen alpha 3 chain protein sequence SEQ ID NO:10.
 KW Human; type IV collagen; anti-angiogenic; angiogenesis; cancer;
 KW benign tumour; rheumatoid arthritis; diabetic retinopathy;
 KW ocular angiogenesis disease; Osler-Weber-Syndrome; telangiectasia;
 KW myocardial angiogenesis; plaque neovascularisation; angiobfibroma;
 KW atherosclerosis; scleroderma; hypertrophic scar; cat scratch disease;
 KW contraception; obesity.
 XX
 OS Homo sapiens.
 XX
 PN WO965940-A1.
 XX
 PD 23-DEC-1999.
 XX
 PF 17-JUN-1999; 99WO-US13737.
 XX
 PR 17-JUN-1998; 98US-0089689.
 PR 25-MAR-1999; 99US-0126175.
 XX
 PA (BETH-) BETH ISRAEL DEACONESS MEDICAL CENT.
 XX
 DR Kalluri R;
 PI
 DR WPI; 2000-097708/08.
 N-PSDB; AAZ257158.
 PT Anti-angiogenic proteins comprising the NC1 domain of the alpha 1, 2 or
 3 chain of type IV collagen used in, e.g. treatment of benign tumors
 and rheumatoid arthritis -
 XX
 PS Claim 32; Fig 16B; 117pp; English.
 XX
 CC The present sequence represents the human type IV collagen alpha 3 chain.
 The present invention describes an isolated protein chosen from the NC1
 domain of the alpha 1, alpha 2 or alpha 3 chains of type IV collagen or
 a fragment, analogue, derivative or mutant, which has anti-angiogenic
 properties. The anti-angiogenic proteins, multimers and chimeras are
 useful for inhibiting angiogenic activity in mammalian tissue,
 especially for treating diseases chosen from angiogenesis-dependent
 cancers, benign tumours, rheumatoid arthritis, diabetic retinopathy,
 psoriasis, ocular angiogenesis diseases, Osler-Webber Syndrome,
 myocardial angiogenesis, plaque neovascularisation, telangiectasia,
 haemophilic joints, angiofibromas, wound granulation, intestinal
 adhesions, atherosclerosis, scleroderma, hypertrrophic scars, cat scratch
 disease, Helicobacter pylori ulcers, dialysis graft vascular access
 stenosis, contraception and obesity. The compositions can be used to
 inhibit a disease characterised by angiogenic activity, in conjunction
 with radiation therapy, chemotherapy or immunotherapy.
 XX
 Sequence 245 AA;
 SQ
 Query Match 100.0%; Score 353; DB 21; Length 245;
 Best Local Similarity 100.0%; Pred. No. 1e-38; 0;
 Matches 64; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 EFRASPFLRCHGRGTCNYYSNSYNSFWLAINPERMFRRKPIPSVKA
 Db 182 efraspflechgrgtnyyssnsysfwlasinpermfrkpiptvka
 QY 61 KKRH 64
 YY 242 kkrh 245
 DE
 RESULT 4
 AAY31993
 ID AAY31993 standard; Protein; 268 AA.
 AC
 XX
 AC AAY31993;
 DT 05-JAN-2000 (first entry)
 XX
 DE Type IV collagen NC1 domain alpha-3 monomer.
 KW Type IV collagen; NC1 domain; non-collagenous domain; human;
 KW angiogenesis; tumour; metastasis; therapy; diabetic retinopathy;
 KW rheumatoid arthritis; retinal neovascularization;
 KW choroidal neovascularization; macular degeneration;
 KW corneal neovascularization; retinopathy of prematurity;
 KW corneal graft rejection; neovascular glaucoma;
 KW retrobulbar fibroplasia; epidemic keratoconjunctivitis;
 KW vitamin A deficiency; contact lens overwear; atopic keratitis;
 KW superior limbic keratitis; pterygium keratitis sicca; sores;
 KW acne rosacea; phylectenulosis; syphilis; Mycobacteria infection;
 KW lipid degeneration; chemical burn; ulcer; herpes simplex infection;
 KW Herpes zoster infection; protozoan infection; Kaposi's sarcoma;

CC	herpes zoster infections; protozoan infections; Kaposi's sarcoma;
CC	marginal keratolysis; trauma; systemic lupus; polyarteritis;
KW	Mooren ulcer; Terrien's marginal degeneration; marginal keratolysis;
KW	Wegener's sarcoidosis; scleritis; Steven's Johnson disease;
KW	radial keratotomy; sickle cell anaemia; sarcoid;
KW	pseudoxanthoma elasticum; Paget's disease; vein occlusion;
KW	artery occlusion; carotid obstructive disease; chronic uveitis;
KW	chronic vitritis; Lyme's disease; Eales disease; Bechets disease;
KW	myopia; optic pit; Stargart' disease; pars planitis;
KW	chronic retinal detachment; hyperviscosity syndrome; toxoplasmosis;
KW	post-laser complication; fibrovascular tissue proliferation;
KW	hemangioma; Osler-Weber-Rendu; AIDS; ocular neovascular disease;
KW	osteoarthritis; chronic inflammation; Crohn's disease;
KW	ulcerative colitis; psoriasis; atherosclerosis; pemphigoid.
XX	XX
OS	Homo sapiens.
OS	Synthetic.
XX	XX
Key	Location/qualifiers
PT	Peptide
FT	/note= "BM40 signal peptide"
FT	Protein
FT	/note= "mature protein"
FT	Peptide
FT	/note= "affinity tag"
FT	Protein
FT	26..268 /note= "NC1 alpha-3 monomer"
XX	XX
PN	W09949885-A2.
XX	XX
PD	07-OCT-1999.
PF	26-MAR-1999; 99WO-US06445.
XX	XX
PR	27-MAR-1998; 98US-0079783.
PR	29-OCT-1998; 98US-0106170.
XX	XX
PA	(UNIV) UNIV KANSAS MEDICAL CENT.
PI	Hudson BG, Sarras MP;
XX	XX
DR	WPI; 1999-601297/51.
N	PSDB; AAZ20091.
XX	XX
PT	Inhibition of angiogenesis with non-collagenous alpha chain monomer useful for treating e.g. tumor growth or metastasis, neovascularisation, etc.
PT	XX
PS	Disclosure; Fig 17c; 56pp; English.
CC	This sequence represents a recombinant type IV collagen non-collagenous (NC1) domain alpha-3 polypeptide composed of a BM40 signal sequence (which is cleaved from the mature protein) to facilitate protein secretion, and a mature protein comprising an affinity tag (facilitates purification and identification of the material) and the alpha-1 chain monomer. The invention provides methods and kits for inhibiting angiogenesis, tumour growth and metastasis, and endothelial cell interaction with the extracellular matrix, each method comprising contacting the tumour or animal tissue with 1 or more isolated type IV collagen NC1 alpha chain monomer(s) selected from the group consisting of alpha-1, alpha-2, alpha-3 and alpha-6 NC1 chain monomers (see Aay31991-96). The monomers can be produced via recombinant protein expression. The polynucleotides and polypeptides are used to treat an angiogenesis-mediated disorder or condition, especially selected from solid and blood-borne tumours, diabetic retinopathy, rheumatoid arthritis, retinal neovascularization, choroidal neovascularization, macular degeneration, corneal neovascularization, retinopathy of prematurity, corneal graft rejection, neovascular glaucoma, retrobulbar fibroplasia, epidemic keratoconjunctivitis, vitamin A deficiency, contact lens overwear, atopic keratitis, superior limbic keratitis, pterygium, keratitis sicca, sognes, acne rosacea, phylectenulosis, syphilis, mycobacteria infections, lipid degeneration, chemical burns, bacterial ulcers, fungal ulcers, herpes simplex infections,
CC	CC
CC	Mooren ulcer; Terrien's marginal degeneration; marginal keratolysis; trauma; systemic lupus; polyarteritis; Wegener's sarcoidosis; scleritis; Steven's Johnson disease; radial keratotomy; sickle cell anaemia; sarcoid; pseudoxanthoma elasticum; Paget's disease; vein occlusion; artery occlusion; carotid obstructive disease; chronic uveitis; chronic vitritis; Lyme's disease; Eales disease; Bechets disease; myopia; optic pit; Stargarts disease; pars planitis; chronic retinal detachment; hyperviscosity syndrome; toxoplasmosis; post-laser complications; abnormal proliferation of fibrovascular tissue; haemangioma; Osler-Weber-Rendu; AIDS; ocular neovascular disease; osteoarthritis; chronic inflammation; Crohn's disease; ulcerative colitis; psoriasis; atherosclerosis; and pemphigoid (all claimed).
CC	CC
CC	herpes zoster infections; protozoan infections; Kaposi's sarcoma; Mooren ulcer; Terrien's marginal degeneration; marginal keratolysis; trauma; systemic lupus; polyarteritis; Wegener's sarcoidosis; scleritis; Steven's Johnson disease; radial keratotomy; sickle cell anaemia; sarcoid; pseudoxanthoma elasticum; Paget's disease; vein occlusion; artery occlusion; carotid obstructive disease; chronic uveitis; chronic vitritis; Lyme's disease; Eales disease; Bechets disease; myopia; optic pit; Stargarts disease; pars planitis; chronic retinal detachment; hyperviscosity syndrome; toxoplasmosis; post-laser complications; abnormal proliferation of fibrovascular tissue; haemangioma; Osler-Weber-Rendu; AIDS; ocular neovascular disease; osteoarthritis; chronic inflammation; Crohn's disease; ulcerative colitis; psoriasis; atherosclerosis; and pemphigoid (all claimed).
XX	Sequence 268 AA.
XX	RESULT 5
ID	AAY9755
XX	Query Match Best Local Similarity 100.0%; Score 353; DB 20; Length 268;
XX	Matches 64; Pred. No. 1; 1e-38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 ERASPLECIEHGRTGTCWVSYNSYSFLNLSPMRKPISTVKAGELEKIIISRCQVM 60
Db	205 efraspflechngrtgtnyyssysfwiaslhpertfrkpkpistvkagelekiisrcqvm 264
QY	61 KRHH 64
Db	265 krhh 268
XX	XX
XX	Query Match Best Local Similarity 100.0%; Score 353; DB 20; Length 268;
XX	Matches 64; Pred. No. 1; 1e-38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
AC	AAY9755;
XX	XX
DR	12-FEB-2001 (first entry)
XX	XX
DE	Human alpha3(IV)NC1 protein sequence.
XX	XX
KW	Type IV collagen alpha chain monomer; human; Inhibitor; angiogenesis; tumour growth; integrin receptor; carcinoma; sarcoma; rhabdomyosarcoma; retinoblastoma; Ewing sarcoma; neuroblastoma; osteosarcoma; leukemia; diabetic retinopathy; rheumatoid arthritis; neovascularisation; muscular degeneration; corneal graft rejection; vitamin A deficiency; atopic keratitis; Mycobacteria infection; chemical burn; sarcoid; kaposi's sarcoma; sickle cell anaemia; carotid obstructive disease; chronic inflammation; psoriasis; therapy; alpha3(IV)NC1.
XX	XX
OS	Homo sapiens.
XX	XX
PN	W020059532-A1.
XX	XX
PD	12-OCT-2000.
XX	XX
PF	31-MAR-2000; 2000WO-US08678.
XX	XX
PR	01-APR-1999; 99US-0127391.
XX	XX
PA	(BIOS-) BIOSTRATUM INC.
XX	XX
PI	Brooks P, Hudson B;
XX	XX
WPI	WPI; 2000-664962/64.
DR	PSDB; AAA80993.
XX	XX
PT	Use of antagonists of specific integrin receptors for inhibiting angiogenesis, tumour growth or metastases, or endothelial cell interactions with the extracellular matrix -
XX	XX
PS	Disclosure; Fig 17c; 78pp; English.
XX	XX

CC This sequence is a human type IV collagen alpha chain monomer, designated alpha(IV)NC1. The invention relates to a method for inhibiting angiogenesis, tumour growth or metastases, or endothelial cell interactions with the extracellular matrix, comprising contacting the cells or tissue with a polypeptide composition containing antagonists of specific integrin receptors. The methods and the antagonists are useful for inhibiting angiogenesis, tumour growth or metastases, or endothelial cell interaction with the extracellular matrix. The antagonists are also useful for treating diseases and conditions with accompanying undesired angiogenesis, e.g. solid and blood borne tumours (e.g. melanomas, carcinomas, sarcomas, rhabdomyosarcoma, retinoblastoma, Ewing sarcoma, neuroblastoma, osteosarcoma or leukaemia). These are also applicable to treating non-tumorigenic diseases and conditions with accompanying undesired angiogenesis, e.g. diabetic retinopathy, rheumatoid arthritis, retinal neovascularisation, choroidal neovascularisation, muscular degeneration, corneal graft rejection, vitamin A deficiency, atopic keratitis, Mycobacteria infections, chemical burns, Kaposi's sarcoma, sickle cell anaemia, sarcoid, carotid obstructive disease, post-laser complications, chronic inflammation or psoriasis.

Sequence 268 AA;

Query Match 100.0%; Score 353; DB 21; Length 268;

Best Local Similarity 100.0%; Pred. No. 1. 1e-38; Matches 64; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EFRASPFLECHGRGTCNYNSYSWLASINPERMKRKP1PSTVKAGELEKTIISRCQVCM 60
Db 205 efraspflechgrgtnyvnsyswlasinpermrkpk1pstvkagelekiisrcqcm 264

Qy 61 KKRH 64

Db 265 kkrh 268

RESULT 6
AAR79164 ID AAR79164 standard; Protein: 218 AA.AC AAR79164;
XX DT 22-DEC-1995 (first entry)

DE Partial sequence of human alpha 3 chain of type IV collagen.

KW Type IV collagen; alpha 3 chain; Alport syndrome; COL4A3 gene.
Homo sapiens.

PN US5424408-A.

XX PD 13-JUN-1995.

XX PF 30-NOV-1990; 900US-0621091.

XX PR 30-NOV-1990; 900US-0621091.

XX PA (UNIV) UNIV KANSAS MEDICAL CENT.

PA (UYYA) UNIV YALE.

PI Hudson BG, Morrison KE, Reenders ST;

XX DR WPI; 1995-262631/34.
N-PSDB; AAQ96291.

XX PT cDNAs encoding human or bovine alpha-3 type 4 collagen peptide(s) - useful for detection and therapeutic removal of antibodies associated with Goodpasture syndrome

PT with Goodpasture syndrome

XX disclosure; columns 7-10; 33PP; English.

XX GC Using the PCR with primers derived from each end of the known 27 AA

CC residue bovine alpha 3 (IV) collagen protein sequence, a 68 bp

CC bovine genomic fragment was amplified. This fragment was then used

CC to a bovine lens cDNA library and a 1.5 kb partial cDNA clone was obtd. (clone KMC15). This encodes 238 residues of the C-terminal non-

CC collagenous (NC1) domain of the alpha 3 (IV) chain. This bovine CC cDNA clone was used to screen a human kidney cDNA library and a CC 2.7 kb human cDNA clone (clone KMC21) was obtd. This clone encodes CC 218 residues of the NC1 domain and a portion of the 3' UTR region CC of the human alpha 3 (IV) chain. The COL4A3 gene localises to CC chromosome 2 and therefore mutations in COL4A3 cannot be CC responsible for Alport syndrome which is X-linked. An isolated CC and substantially pure nt. having the sequence in AAQ96291 is claimed.

Sequence 218 AA;

Query Match 95.8%; Score 338; DB 16; Length 218;

Best Local Similarity 96.9%; Pred. No. 8.8e-37; Matches 62; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 EFRASPFLECHGRGTCNYNSYSWLASINPERMKRKP1PSTVKAGELEKTIISRCQVCM 60
Db 155 efpaspflechgrgtnyvnsyswlasinpermrkpk1pstvkagelekiisrcqcm 214

Qy 61 KKRH 64
Db 215 kkrh 218

RESULT 7
AAR79163 ID AAR79163 standard; Protein: 471 AA.

XX AC AAR79163;
XX DT 22-DEC-1995 (first entry)

DE Partial sequence of bovine alpha 3 chain of type IV collagen.

XX KW Type IV collagen; alpha 3 chain.

XX OS Bos taurus.

XX PN US5424408-A.

XX PD 13-JUN-1995.

XX PF 30-NOV-1990; 900US-0621091.

XX PR 30-NOV-1990; 900US-0621091.

XX PA (UNIV) UNIV KANSAS MEDICAL CENT.

PA (UYYA) UNIV YALE.

XX PI Hudson BG, Morrison KE, Reenders ST;

XX DR WPI; 1995-262631/34.
N-PSDB; AAQ96290.

XX PT cDNAs encoding human or bovine alpha-3 type 4 collagen peptide(s) - useful for detection and therapeutic removal of antibodies associated with Goodpasture syndrome

XX PS Disclosure; Columns 5-8; 33PP; English.

XX CC Using the PCR with primers derived from each end of the known 27 AA

CC residue bovine alpha 3 (IV) collagen protein sequence, a 68 bp

CC bovine genomic fragment was amplified. This fragment was then used

CC to a bovine lens cDNA library and a 1.5 kb partial cDNA clone was obtd. (clone KMC15). This encodes 238 residues of the triple helical

THIS PAGE BLANK (USPTO)

Gencore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on:

March 6, 2002, 06:55:22 ; Search time 26.9 Seconds

(without alignments)
53.539 Million cell updates/sec

Title: US-09-543-371-10-COPY_181_244

Perfect score: 353
Sequence: 1 EFRASPFLECHGRGTCNYYS..... KAGELEKIIISRCQVCMKKRH 64

Scoring table: BLOSUM62

Gapext 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

number of hits satisfying chosen parameters: 6

Minimum DB seq length: 0
Maximum DB seq length: 200000000Post-processing: Minimum Match 90%
Maximum Match 100%

Database : Issued Patents AA:*

Listing first 1000 summaries

Score: 1

/cen2_6/podata/2/1aa/5A_COMBO.pep:*

2: /cg2_5/podata/2/1aa/5B_COMBO.pep:*

3: /cg2_6/podata/2/1aa/6A_COMBO.pep:*

4: /cen2_6/podata/2/1aa/6B_COMBO.pep:*

5: /cen2_5/podata/2/1aa/PCUS_COMBO.pep:*

6: /cen2_6/podata/2/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	353	100.0	218	US-09-399-889-25
2	353	100.0	218	US-09-399-889-25
3	333	100.0	218	US-09-439-897-4
4	333	91.5	471	US-09-399-889-24
5	323	91.5	471	US-09-167-364-24
6	323	91.5	471	US-09-439-889-2

ALIGNMENTS

RESULT 1
US-09-399-889-25
; Sequence 25, Application US/08399889B
; Patient No. 5973120
; GENERAL INFORMATION:
; APPLICANT: Reeder, Stephen T
; APPLICANT: Morrison, Karen E
; TITLE OF INVENTION: Alpha-3 Chain Type IV Collagen Polypeptides
; FILE REFERENCE: 951263B
; CURRENT APPLICATION NUMBER: US/09/167_364
; CURRENT FILING DATE: 1998-10-07
; EARLIER APPLICATION NUMBER: 08/399889
; EARLIER FILING DATE: 1995-03-07
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO: 25
; LENGTH: 218
; TYPE: PRT
; ORGANISM: Human
; US-09-167-364-25

RESULT 2
US-09-167-364-25
; Sequence 25, Application US/09167364
; Patient No. 600780
; GENERAL INFORMATION:
; APPLICANT: Reeder, Stephen T
; APPLICANT: Morrison, Karen E
; TITLE OF INVENTION: Alpha-3 Chain Type IV Collagen Polypeptides
; FILE REFERENCE: 951263B
; CURRENT APPLICATION NUMBER: US/09/167_364
; CURRENT FILING DATE: 1998-10-07
; EARLIER APPLICATION NUMBER: 08/399889
; EARLIER FILING DATE: 1995-03-07
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO: 25
; LENGTH: 218
; TYPE: PRT
; ORGANISM: Human
; US-09-167-364-25

RESULT 3
US-09-439-897-4
; Sequence 4, Application US/09439897
; Patient No. 6277558
; GENERAL INFORMATION:
; APPLICANT: Hudson, Billy G
; TITLE OF INVENTION: Alpha-3 Chain Type IV Collagen Polypeptides
; FILE REFERENCE: 95-1263-C
; CURRENT APPLICATION NUMBER: US/09/439,897
; CURRENT FILING DATE: 1999-11-12
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO: 4
; LENGTH: 218
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-439-897-4

Query Match 100.0%; Score 353; DB 4; Length 218;
 Best Local Similarity 100.0%; Pred. No. 1.6e-39; Mismatches 0;
 Matches 64; Conservative 0; Indels 0; Gaps 0;

QY 1 EFRASPFLECHGRGTCNYYSNSYSFWLASNPERMFRKPKIPSTVKAGELEKITSRCQVM 60
 Db 155 EFRASPFLECHGRGTCNYYSNSYSFWLASNPERMFRKPKIPSTVKAGELEKITSRCQVM 214

QY 61 KKRH 64
 Db 215 KRRH 218

RESULT 4
 US-08-399-889-24
 Sequence 24, Application US/08399889B
 Patent No. 597120

GENERAL INFORMATION:
 APPLICANT: Reeders, Stephen T
 APPLICANT: Morrison, Karen E
 APPLICANT: Hudson, Billy G
 TITLE OF INVENTION: Alpha 3 Chain Type IV Collagen Polypeptides

CURRENT APPLICATION NUMBER: US/08/399, 889B
 CURRENT FILING DATE: 1995-03-07
 EARLIER APPLICATION NUMBER: 07/621091

EARLIER FILING DATE: 1990-11-30

NUMBER OF SEQ ID NOS: 25
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 24
 LENGTH: 471
 TYPE: PRT
 ORGANISM: Calf

US-08-399-889-24

Query Match 91.5%; Score 323; DB 2; Length 471;
 Best Local Similarity 92.1%; Pred. No. 4.1e-35; Mismatches 3;
 Matches 58; Conservative 3; Indels 0; Gaps 0;

QY 1 EFRASPFLECHGRGTCNYYSNSYSFWLASNPERMFRKPKIPSTVKAGELEKITSRCQVM 60
 Db 408 EFRASPFLECHGRGTCNYYSNSYSFWLASLPKMRKPKIPSTVKAGELENTISRCQVM 467

QY 61 KKR 63
 Db 468 KMR 470

RESULT 5
 US 09-167-364-24
 Sequence 24, Application US/09167364
 Patent No. 6007980

GENERAL INFORMATION:
 APPLICANT: Reeders, Stephen T
 APPLICANT: Morrison, Karen E
 APPLICANT: Hudson, Billy G
 TITLE OF INVENTION: Alpha-3 Chain Type IV Collagen Polypeptides

FILE REFERENCE: 951263B
 CURRENT APPLICATION NUMBER: US/09/167, 364
 CURRENT FILING DATE: 1998-10-07
 EARLIER APPLICATION NUMBER: 08/399889
 EARLIER FILING DATE: 1995-03-07
 NUMBER OF SEQ ID NOS: 25
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 24
 LENGTH: 471
 TYPE: PRT
 ORGANISM: Calf

US-08-167-364-24

Query Match 91.5%; Score 323; DB 4; Length 471;
 Best Local Similarity 92.1%; Pred. No. 4.1e-35; Mismatches 2;
 Matches 58; Conservative 3; Indels 0; Gaps 0;

QY 1 EFRASPFLECHGRGTCNYYSNSYSFWLASNPERMFRKPKIPSTVKAGELEKITSRCQVM 60
 Db 408 EFRASPFLECHGRGTCNYYSNSYSFWLASLPKMRKPKIPSTVKAGELENTISRCQVM 467

QY 61 KKR 63
 Db 468 KMR 470

Search completed: March 6, 2002, 06:55:22
 Job time: 73 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 6, 2002, 06:54:09 ; Search time 31.16 Seconds

(without alignments)
466.924 Million cell updates/sec

Title:

US-09-543-371-10_COPY_54_244
Perfect score: 1052
Sequence: 1 QRMHGQDGLTGSLQRFTT..... KAGELEKIIISRQCVCMKKRRH 191

Scoring table:

BLOSUM62
Gap0 10.0 , Gapext 0.5

Searched:

219241 seqs, 76174552 residues

Tr. number of hits satisfying chosen parameters: 3

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 90%
Maximum Match 100%
Listing first 1000 summaries

Database : PIR:68:
1: pir1:
2: pir2:
3: pir3:
4: pir4:
5:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match Length	DB ID	Description
1	1052	100.0	1670	1 CGHU3B	collagen alpha 3(I)
2	993	94.4	246	2 I48302	collagen alpha 3(I)
3	988	93.9	471	2 A39024	collagen alpha 3(I)

ALIGNMENTS

RESULT	1
CGHU3B	N:Alternate names: Goodpasture antigen; procollagen alpha 3(IV) chain precursor, long splice form - human collagen alpha 3(IV) chain precursor, long splice form - human collagen alpha 3(IV) chain long splice form C:Species: Homo sapiens (man) C:Date: 28-Oct-1994 #sequence_revision 03-Oct-1995 #text_change 22-Jun-1999 C:Accession: A54763; A43928; A44043; A45971; A39786 R:Mariyama, M.; Leinonen, A.; Mochizuki, T.; Tryggvason, K.; Reeder, S.T. J. Biol. Chem. 268, 23013-23017, 1994 A:Title: Complete primary structure of the human alpha3(IV) collagen chain. Coexpression of the molecular cloning of the human Goodpasture antigen demonstrates it to be the antigenic determinant. A:Reference number: A54763; MUID:94364994 A:Molecule type: mRNA A:Residues: 1-1670 <MAR> A:Cross-references: GB:80031; NID:9577563; PID:9577564 A:Experimental source: kidney R:Turner, N.; Mason, P.J.; Brown, R.; Fox, M.; Povey, S.; Rees, A.; Pusey, C.D. J. Clin. Invest. 89, 592-601, 1992 A:Title: Molecular cloning of the human Goodpasture antigen demonstrates it to be the antigenic determinant. A:Reference number: A43928; MUID:92147878 A:Accession: A43928

A: Molecule type: mRNA
A: Residues: 1331-1524, 'I', 1526-1670 <TUR>
A: Cross-references: GB:M81379
A: Experimental source: kidney R:Quinones, S.; Bernal, D.; Garcia-Sogo, M.; Elena, S.F.; Saus, J. J. Biol. Chem. 267, 19780-19784, 1992
A: Title: Exon/intron structure of the human alpha 3(IV) gene encompassing the Goodpasture antigen. A: Reference number: A44043; MUID:93015826
A: Molecule type: DNA; mRNA
A: Residues: 1386-1670 <QUIT>
A: Cross-references: GB:M9293; NID:9177895; PID:AA21610.1; PID:9177896
A: Note: sequence extracted from NCBI backbone (NCBIP:133363); sequence incorrectly identified.
R:Quinones, S.; Bernal, D.; Garcia-Sogo, M.; Elena, S.F.; Saus, J. J. Biol. Chem. 269, 17358, 1994
A: Reference number: A44738; MUID:94274734
A: Reference number: A45971; MUID:93280184
A: Contents: annotation; erratum; correction to intronic sequence in A44043
A: Status: nucleic acid sequence not shown
A: Molecule type: mRNA
A: Residues: 1427-1444 <BER>
A: Note: sequence extracted from NCBI backbone (NCBIP:115597)
R: Morrison, K.E.; Mariyama, M.; Yang-Feng, T.L.; Reeders, S.T.
Am., J. Hum. Genet. 49, 545-554, 1991
A: Title: Sequence and localization of a partial cDNA encoding the human alpha3 chain
A: Reference number: A39786; MUID:91353570
A: Accession: A39786
A: Gene: GDB:COL4A3
A: Molecule type: mRNA
A: Residues: 1453-1593, 'A', 1595-1670 <MOR>
A: Cross-references: GB:S55790; NID:9234418; PID:AB19637.1; PID:9234419
C:Comment: In Goodpasture's syndrome, an autoimmune response develops against an epitope containing prolines and lysines at the third position of the tripeptide repeating unit and subsequently O-glycosylated.
C:Genetics:

A: Cross-references: GDB:128351; OMIM:120070
A: Map Position: Zq36-2037
A: Note: the alpha 3(IV) and alpha 4(IV) chain genes are encoded on opposite strands w.r.t. the alpha 1(IV) chain gene.
C:Complex: This minor type IV collagen is thought to form a heterotrimer of two alpha monomer trimer amino-terminal domains (with disulfide and desmosine cross-links), dimeric associations in the interrupted helical domain (with disulfide and desmosine cross-links).
C:Function:
A: Description: minor structural component of extracellular basement membrane in kidney
C:Superfamily: collagen alpha 1(IV) chain
C:Keywords: alternative splicing; basement membrane; cell binding; coiled coil; extra F:1-28/Domain: signal sequence #status predicted <SG>
F:29-1670/Domain: collagen alpha 3(IV) chain long splice form #status predicted <MA>
F:29-42/Domain: amino-terminal nonhelical, NHI <NHI>
F:43-1438/Region: interrupted helical F:791-93/Region: cell attachment (R-G-D) motif F:966-998/Region: cell attachment (R-G-D) motif F:1154-1156/Region: cell attachment (R-G-D) motif F:1306-1308/Region: cell attachment (R-G-D) motif F:1345-1347/Region: cell attachment (R-G-D) motif F:1432-1434/Region: cell attachment (R-G-D) motif F:1439-1670/Domain: carboxyl-terminal nonhelical, NCI <NCI>
F:1451-1551/Domain: collagen IV carboxyl-terminal repeat <CTL>
F:1561-1665/Domain: collagen IV carboxyl-terminal repeat <CT2>
F:253/Binding Site: carbohydrate (Asn) (covalent) #status predicted F:1505-1548,1493-1551/Disulfide bonds: #status predicted F:1460-1511,1616-1622/Disulfide bonds: #status predicted F:1570-1662,1604-1665/Disulfide bonds: (or 1570-1665, 1604-1662) #status predicted

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

RP
SEQUENCE OF 1386-1670 FROM N.A.; AND PARTIAL SEQUENCE.
RX
MEDLINE=93015826; PubMed=1400291;

OM protein - protein search, using sw model

Run on : March 6, 2002, 06:55:24 ; Search time 18.75 Seconds
 (w/o alignments)
 373,492 Million cell updates/sec

Title: US-09-543-371-10_COPY_54_244
Perfect score: 1052

Sequence: 1 ORAHGQDLGLTGSCLQRFTT KAGELEKIIISRCQVCMKKRH 1911

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 2

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 90%, Maximum Match 100%, Listing first 100

Database : SwissProt_39:*

Pred. No. is the number of results predicted score greater than or equal to the score

and is derived by analysis of the total score distribution.

SUMMARIES

NO.	Score	Match	Length	DB	ID	Description
1	1052	100 0	1670	1	CIA4 HITMAN	001055 home canon

2 - 988 93.9 471 1 CA34_BOVIN . 028084 bos taurus

ALIGNMENTS

RESULT 1

C:\HUMAN FA34_HUMAN STANDARD; PRT; 1670 AA
ACC# 001955:

DT 01-OCT-1996 (Rel. 34, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)

DT 2014-08-20 {ReI: 40', Last annotation update: 2011-05-10
DE COLLAGEN ALPHA 3 (IV) CHAIN PRECURSOR (GOODP
GN COL4A3.

OS
OC
Homo sapiens (Human)
Eukaryota; Metazoa; Chordata; Craniata; Ver-

Mammalia; Eutheria; Primates; Catarrhini; H
OC_OX_RN [I]
Mammalia; Eutheria; Primates; Catarrhini; H
OC_OX_RN [I]

RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RT 1970-01-01 1970-01-01

RT Coexpression of the alpha 3(IV) and alpha 4(IV) collagen chains in RT human tissues." J CO-20013 33017/1004
BY RICARDO CHAVES

J. BOL. CHIN. 203:23013-23017 (1994).
[2] REVISIONS.

RA Leinonen A.;
RL Submitted (OCT-1998) to the EMBL/GenBank/DDBJ databases.

CC MESHWORK TOGETHER WITH LAMININS, PROTEOGLYCANS AND ENACTIN/ NIDGEN. CC SUBUNITS: THERE ARE SIX TYPE IV COLLAGEN ISOFORMS, ALPHA 1(IV)- ALPHA 6(IV), EACH OF WHICH CAN FORM A TRIPLE HELIX STRUCTURE WITH 2 OTHER CHAINS TO GENERATE TYPE IV COLLAGEN NETWORK.

CC -!- SUBCELLULAR LOCATION: CELL SURFACE (POTENTIAL).

CC -!- ALTERNATIVE PRODUCTS: AT LEAST 3 ISOFORMS; 1 (SHOWN HERE), 2/V AND 3/L5, ARE PRODUCED BY ALTERNATIVE SPLICING. THEY DIFFER IN THEIR C-TERMINAL NCI DOMAINS.

CC -!- TISSUE SPECIFICITY: ALPHA 3 AND ALPHA 4 TYPE IV COLLAGENS ARE COLOCALIZED AND PRESENT ONLY IN BASEMENT MEMBRANES OF KIDNEY, EYE, COCHLEA, LONG AND BRAIN.

CC -!- DOMAIN: ALPHA CHAINS OF TYPE IV COLLAGEN HAVE A NONCOLLAGENOUS DOMAIN (NCI) AT THEIR C-TERMINUS. FREQUENT INTERRUPTIONS OF THE G-X-Y REPEATS IN THE LONG CENTRAL 'TRIPLE HELICAL DOMAIN' (WHICH MAY CAUSE FLEXIBILITY IN THE TRIPLE HELIX), AND A SHORT N-TERMINAL TRIPLE-HELICAL 7S DOMAIN.

CC -!- PTM: PROLINES AT THE THIRD POSITION OF THE TRIPETIDE REPEATING UNIT (G-X-Y) ARE HYDROXYLATED IN SOME OR ALL OF THE CHAINS.

CC -!- PTM: THE ALTERNATIVE SPliced FORM V CONTAINS AN ADDITIONAL N-LINKED GLYCOSYLATION SITE.

CC -!- PTM: TYPE IV COLLAGENS CONTAIN NUMEROUS CYSTEINE RESIDUES WHICH ARE INVOLVED IN INTER- AND INTRAMOLECULAR DISULFIDE BONDING. 12 OF THESE, LOCATED IN THE NCI DOMAIN, ARE CONSERVED IN ALL KNOWN TYPE IV COLLAGENS.

CC -!- DISEASE: ANTIBODIES AGAINST THE NCI DOMAIN OF ALPHA 3(IV) MEDIATE THE HUMAN AUTOIMMUNE DISEASE, GOODPASTURE SYNDROME, WHICH IS CHARACTERIZED BY HEMATURIA AND PULMONARY HEMORRHAGE.

CC -!- DISEASE: DEFECTS IN COL4A3 ARE ASSOCIATED WITH THE TYPE I AUTOSOMAL RECESSIVE FORM OF ALPORT SYNDROME, AN HEREDITARY HEMATURIA AND DEAFNESS. THE RECESSIVE FORM OCCURS EQUALLY BETWEEN MALES AND FEMALES.

CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC DR X80031; CA056335; 1; -.

DR EMBL; M22993; AA21610; 1; -.

DR EMBL; M55790; AAB19637; 1; -.

DR EMBL; M81379; AA51556; 1; -.

DR EMBL; U08650; AA52044; 1; -.

DR EMBL; U02519; AA18942; 1; -.

DR EMBL; U02520; AA18943; 1; -.

DR EMBL; AB008495; BAB25064; 1; -.

MIM; 120070; -.

DR MIM; 203780; -.

DR MIM; 233450; -.

DR InterPro; IPR01442; C4.

DR InterPro; IPR000087; Collagen.

DR Pfam; PF01413; C4; 2.

DR Pfam; PF01391; Collagen; 21.

DR ProDom; PD005923; C4; 2.

DR SMART; SM00111; C4; 2.

KW Extracellular matrix; Connective tissue; Repeat; Hydroxylatation; Glycoprotein; Basement membrane; Collagen; Signal; Cell adhesion; Alternative splicing; Polymorphism; Phosphorylation; Disease mutation; KW Alport syndrome.

FT SIGNAL 1 28 POTENTIAL.

FT CHAIN 29 1670 COLLAGEN ALPHA 3 (IV) CHAIN.

FT DOMAIN 29 42 75 DOMAIN.

FT DOMAIN 43 1438 TRIPLE-HELICAL REGION.

FT DOMAIN 1439 1670 NONHELICAL REGION (NCI). (GOODPASTURE ANTIGEN) (BY SIMILARITY).

FT DOMAIN 1427 1444 EPITOPE (RECOGNIZED BY GOODPASTURE ANTIBODIES).

FT SITE 1426 1427 CLEAVAGE (BY COLLAGENASE) (BY SIMILARITY).

FT SITE 791 793 CELL ATTACHMENT SITE (POTENTIAL).

FT SITE 996 998 CELL ATTACHMENT SITE (POTENTIAL).

FT SITE 1154 1156 CELL ATTACHMENT SITE (POTENTIAL).

FT SITE 1306 1308 CELL ATTACHMENT SITE (POTENTIAL).

FT SITE 1345 1347 CELL ATTACHMENT SITE (POTENTIAL).

FT SITE 1434 1434 CELL ATTACHMENT SITE (POTENTIAL).

FT CARBOHYD 253 253 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT MOD_RES 1435 1435 PHOSPHORYLATION (BY SIMILARITY).

FT MOD_RES 1437 1437 PROPHYLATION (BY SIMILARITY).

FT DISULFID 1460 1551 OR 1548 (BY SIMILARITY).

FT DISULFID 1493 1548 OR 1551 (BY SIMILARITY).

FT DISULFID 1570 1665 BY SIMILARITY.

FT DISULFID 1604 1662 OR 1665 (BY SIMILARITY).

FT DISULFID 1616 1622 BY SIMILARITY.

FT VARSPPLIC 1586 1670 FUSING SECTGOALASPOSCLEEFRRASPFPLECHRGRTGNYYS NSISFLASLNPERMRKKPPIPSTVAGELK1ISRCQNCVKM KRH-> KAYSINCNSPERMRKKPPIPSTVAGELK1ISRCQNCVKM ELVFETLKNKVMVEHAVI (IN ISFORM 2).

FT VARSPPLIC 1488 1670 GTGCGCLOFTTMPLFCNVNDCNFNSRNDYNSWLTAPMLPNNMAPITGR 60 MPMMAPITGRALAEPYYSRCTCEGPAIAVHSQTFDPCPGHWISLWKGSFIMTSAGSEBTGQALA 1539 CPGHWSLWKGSFIMTSAGSEBTGQALA SPFECLEHGRTGTCYNSISFLASLNPERMRKKPPIPSTVAGELK1 1540 AGELK1ISRCQNCVKM KRH-> DALFKVLRSP (IN ISFORM 3).

FT VARIANT 43 43 G-> R.

FT VARIANT 162 162 /FTId=VAR_011202.

FT VARIANT 297 297 G-> E.

Query Match 100.0%; Score 1052; DB 1; Length 1670; Best Local Similarity 100.0%; Pred. No. 3; 5e-96; Matches 191; Conservative 0; Mismatches 0; Indels 0; Gaps 0; Query OY 1 QRAHQDQDITGSGCQLQRTFTMPFLFCNNDVCNFNASRNDYNSWLTALMPNNMAPITGR 60 Db 1480 QRAHQDQDITGSGCQLQRTFTMPFLFCNNDVCNFNASRNDYNSWLTALMPNNMAPITGR 1539 OY 61 ALEPYTISRCTVCEGPAIAVHSQTFDPCPGHWISLWKGSFIMTSAGSEBTGQALA 120 Db 1540 ALEPYTISRCTVCEGPAIAVHSQTFDPCPGHWISLWKGSFIMTSAGSEBTGQALA 1599 OY 121 SPGSCLEEFRRASPFPLECHRGRTGTCNYYSNSYSFWLASLNPERMRKKPPIPSTVAGELK1 180 Db 1600 SPGSCLEEFRRASPFPLECHRGRTGTCNYYSNSYSFWLASLNPERMRKKPPIPSTVAGELK1 1659 OY 181 SRCQVCMKKH 191 Db 1660 SRCQVCMKKH 1670

RESULT 2

CA34_BOVIN CA34_BOVIN STANDARD; PRT: 471 AA.

TD CA34_BOVIN STANDARD;

AC 028084;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 01-NOV-1997 (Rel. 35, Last annotation update)

DE COLLAGEN ALPHA 3(IV) CHAIN (FRAGMENT).

GN COL4A3.

OS Bos taurus (Bovine).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Buteraria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;

OC Bovidae; Bovinae; Bos.

OX NCBI_TaxID=9913;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Lens;

RX MEDLINE=91093145; PubMed=1985905;

RA Morrison K.E., Germino G.G., Reeder S.T.;

RT "use of the polymerase chain reaction to clone and sequence a cDNA encoding the bovine alpha 3 chain of type IV collagen";

J. BIOL. CHEM. 266:34-39(1991).
 RP RL
 RP SEQUENCE OF 227-258.
 TISSUE=Kidney;
 MEDLINE=90202779; PubMed=2318822;
 RA Gunwar S.; Saus J.; Noecken M.E.; Hudson B.G.;
 "Glomerular basement membrane. Identification of a fourth chain,
 alpha 4, of type IV collagen;"
 RL J. Biol. Chem. 265:5466-5469(1990).
 RT [31]
 SEQUENCE OF 227-254.
 RX MEDLINE=68330844; PubMed=3417661;
 RA Saus J.; Wieslander J.; Langeveld J.P.M.; Quinones S.; Hudson B.G.;
 "Identification of the Goodpasture antigen as the alpha 3(IV) chain
 of collagen IV;"
 RL J. Biol. Chem. 263:13374-13380(1988).
 RN [4]
 RP SEQUENCE OF 227-244.
 RX MEDLINE=8722419; PubMed=2438283;
 RA Burkowski R.J.; Langeveld J.P.M.; Wieslander J.; Hamilton J.,
 Hudson B.G.;
 RL Localization of the Goodpasture epitope to a novel chain of basement
 membrane collagen;"
 RL J. Biol. Chem. 262:7874-7877(1987).
 CC -1- FUNCTION: TYPE IV COLLAGEN IS THE MAJOR STRUCTURAL COMPONENT OF
 GLOMERULAR BASEMENT MEMBRANES (GBM), FORMING A "CHICKEN-WIRE"
 MESHWORK TOGETHER WITH LAMININS, PROTEOGLYCANS AND ENACTIN/
 NIDOCIN.
 CC -1- SUBUNIT: THERE ARE SIX TYPE IV COLLAGEN ISOFORMS, ALPHA 1(IV)-
 ALPHA 6(IV), EACH OF WHICH CAN FORM A TRIPLE HELIX STRUCTURE
 WITH 2 OTHER CHAINS TO GENERATE TYPE IV COLLAGEN NETWORK.
 CC -1- SUBCELLULAR LOCATION: CELL SURFACE (POTENTIAL).
 CC -1- DOMAIN: ALPHA CHAINS OF TYPE IV COLLAGEN HAVE A NONCOLLAGENOUS
 DOMAIN (NC1) AT THEIR C-TERMINUS, FREQUENT INTERRUPTIONS OF THE
 G-X-Y REPEATS IN THE LONG CENTRAL TRIPLE-HELICAL DOMAIN (WHICH MAY
 CAUSE FLEXIBILITY IN THE TRIPLE HELIX), AND A SHORT N-TERMINAL
 TRIPLE HELICAL 7S DOMAIN.
 CC -1- PTM: PROLINE AT THE THIRD POSITION OF THE TRIPÉPTIDE REPEATING
 UNIT (G-X-Y) ARE HYDROXYLATED IN SOME OR ALL OF THE CHAINS.
 CC -1- PTM: TYPE IV COLLAGENS CONTAIN NUMEROUS CYSTEINE RESIDUES WHICH
 ARE INVOLVED IN INTR- AND INTRAMOLECULAR DISULFIDE BONDING. 12 OF
 THESE, LOCATED IN THE NC1 DOMAIN, ARE CONSERVED IN ALL KNOWN TYPE
 IV COLLAGENS.
 CC -1- SIMILARITY: TO OTHER TYPE IV COLLAGENS.
 CC ---
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC ---
 DR EMDU: M63139; AAA2708 1;
 DR InterPro: IPR001442; C4;
 DR Pfam: PF01413; C4; 2.
 DR ProDom: PD005923; C4; 2.
 DR SMART: SM00111; C4; 2.
 DR InterPro: IPR001442; C4;
 DR InterPro: IPR000087; Collagen.
 DR Pfam: PF01391; Collagen; 4.
 DR ProDom: PD005923; C4; 2.
 KW Extracellular matrix; connective tissue; Repeat; Hydroxylation;
 KW Glycoprotein; Basement membrane; Collagen; Cell adhesion.
 FT NON_TER 1 1
 FT DOMAIN <1 238 TRIPLE-HELICAL REGION.
 FT DOMAIN 239 471 NONHELICAL REGION (NC1).
 SITE 106 108 CELL ATTACHMENT SITE (POTENTIAL).
 FT MOD_RES 232 232 HYDROXYLATION.
 FT MOD_RES 238 238 HYDROXYLATION.
 FT DISULFID 261 352 OR 349 (BY SIMILARITY).
 FT DISULFID 294 349 OR 352 (BY SIMILARITY).
 FT DISULFID 306 312 BY SIMILARITY.
 FT DISULFID 371 466 OR 463 (BY SIMILARITY).
 FT DISULFID 405 463 OR 466 (BY SIMILARITY).

Search completed: March 6, 2002, 07:05:08
Total time: 584 sec

THIS PAGE BLANK (USPTO)